2004

Toeplitz Operators on Locally Compact Abelian Groups

David Gaebler
Harvey Mudd College

Toeplitz Operators on Locally Compact Abelian Groups

David J. Gaebler

Henry A. Krieger, Advisor

Michael R. Raugh, Reader

May, 2004

HARVEY MUDD COLLEGE

Department of Mathematics

Abstract

Given a function (more generally, a measure) on a locally compact Abelian group, one can define the Toeplitz operators as certain integral transforms of functions on the dual group, where the kernel is the Fourier transform of the original function or measure. In the case of the unit circle, this corresponds to forming a matrix out of the Fourier coefficients in a particular way. We will study the asymptotic eigenvalue distributions of these Toeplitz operators.

Contents

Chapter 1

The General Problem

If G is a locally compact Abelian group, Γ its dual group, and $f \in L^1(G)$, then the Toeplitz operator generated by f is the linear operator T_f on $L^2(\Gamma)$ given by

$$(T_f\phi)(\gamma) = \int_\Gamma \hat{f}(\gamma - \tau)\phi(\tau)\,d\tau$$

for $\phi \in L^2(\Gamma)$. (See Appendix A for details of Fourier analysis on locally compact Abelian groups.)

We are interested in studying the spectrum of this operator. We restrict our attention to real f, for which T_f is Hermitian. In general, the spectrum will be continuous. However, if it happens that G is discrete, so that Γ is compact, then T_f is a Hilbert-Schmidt operator and is therefore a compact operator. This means that the spectrum will be discrete with 0 as the only possible limit point. This allows us to study the distribution of eigenvalues by considering the number of eigenvalues greater than some real number, or the number in some interval on the real line (which interval should exclude zero to guarantee that it contains finitely many eigenvalues).

If Γ is not compact, we form "finite Toeplitz operators" by integrating over compact subsets of Γ. Intuitively, if we can find a sequence of such compact subsets which expands to fill all of Γ in some nice way, the limiting distribution of the corresponding finite Toeplitz operators should tell us something meaningful about f. The following theorem from [Krieger] confirms that this is the case:

Theorem 1. *Let G be a non-discrete LCA group, Γ its non-compact dual group, and m and ν Haar measures on G and Γ, normalized so that the Fourier inversion theorem holds. Let $f \in L^1(G)$, and for a Borel set $W \subset \Gamma$ with compact closure, define the operator F_W on $L^2(W)$ by $(T_W\phi)(\gamma) = \int_W \hat{f}(\gamma - \tau)\phi(\tau)\,d\tau$. If Γ*

is compactly generated, there exists a sequence $\{W_n\}$ of Borel sets such that if $\lambda_j^{(n)}$, $j = 1, 2, \ldots$, are the eigenvalues of T_{W_n}, and $[a, b] \subset \mathbb{R}$ is an interval not containing zero for which

$$m(f^{-1}(\{a\})) = m(f^{-1}(\{b\})) = 0,$$

then

$$\lim_{n\to\infty} \frac{\#\text{ of } \lambda_j^{(n)} \text{ in } [a,b]}{\nu(W_n)} = m(f^{-1}([a,b])).$$

We will explore generalizations of this theorem. Let $\mu \in M(G)$ be a regular finite complex measure on G. We can define the Toeplitz operators generated by μ in a manner entirely analogous to the above by using the Fourier-Stieltjes transform $\hat{\mu}$. Thus, we define

$$(T_\mu \phi)(\gamma) = \int_\Gamma \hat{\mu}(\gamma - \tau)\phi(\tau)\, d\tau$$

and

$$(T_{W_n}\phi)(\gamma) = \int_{W_n} \hat{\mu}(\gamma - \tau)\phi(\tau)\, d\tau.$$

Note that this contains the previously defined Toeplitz operators as a special case, viz. absolutely continuous measures.

To avoid excessive notational clutter, we will refrain from indicating both the measure and the compact subset of Γ used to generate a given operator (we could call it something like T_{μ,W_n} if we really wanted to); this will not present any difficulty as it will always be clear from the context what the measure is.

It is natural to wonder what sort of distribution limits will hold for the Toeplitz operators generated by a measure. We already know the answer for absolutely continuous measures; how will the answer change due to the addition of a singular part? A possibility that comes to mind, and that would be desirable to establish, is that the singular part has no effect at all on the asymptotic eigenvalue distribution. It has been proven in [Krieger] that the addition of a discrete part will not affect the limiting distribution; however, the corresponding question for continuous singular measures is still open. This thesis will attempt to answer it in certain special cases.

Chapter 2

The Unit Circle

In this chapter we will consider the simplest case of a Toeplitz operator, for which the relevant groups are the unit circle and the integers. In this case the Toeplitz operator may be represented by a matrix, which allows powerful tools from linear algebra to be brought to bear on the calculation of the eigenvalues.

Throughout the chapter, we use T to denote the unit circle, parametrized by $\theta \in [-\pi, \pi)$. We use $d\theta$ for ordinary Lebesgue measure, and dm for normalized Lebesgue measure (i.e. $dm = \frac{1}{2\pi} d\theta$); this is just the standard normalization for Haar measure (Appendix A.2).

Definition 1. *Let $f : T \to \mathbb{R}$ with $f \in L^1(T)$. Then the nth Toeplitz matrix generated by f is the matrix T_n with entries*

$$(T_n)_{i,j} = c_{i-j} \qquad i, j = 0, \ldots, n$$

where $c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\theta} f(\theta)\, d\theta = \int e^{-in\theta} f\, dm$ is the nth Fourier coefficient of f.

Note that since f is real-valued, $c_{-k} = \overline{c_k}$ so that T_n is Hermitian. In particular, the spectrum of T_n is real. Also observe that with this notation, T_n is an $(n+1)$ by $(n+1)$ matrix.

Each T_n has an associated quadratic form, which we naturally call the nth Toeplitz form generated by f and denote by Q_n:

$$Q_n(\mathbf{u}) = \mathbf{u}^* \, T_n \, \mathbf{u}, \qquad \mathbf{u} \in \mathbb{C}^{n+1}$$

It is easy to check that

$$Q_n(\mathbf{u}) = \int_{-\pi}^{\pi} |u_0 + u_1 e^{i\theta} + \cdots + u_n e^{in\theta}|^2 \, f(\theta)\, dm.$$

Thus, one way to view the Toeplitz form is as follows: Given a complex vector, it returns the integrated square length of the corresponding trigonometric polynomial, where the integral is with respect to the weighting function f. In other words, it returns the square of the $L^2(\mu)$ norm, where $\mu = f\,dm$.

Since $\sum_{k=0}^{n} |u_k|^2 = \int_{-\pi}^{\pi} |u_0 + u_1 e^{i\theta} + \cdots + u_n e^{in\theta}|^2 \, dm$, we see that for essentially bounded f the eigenvalues of T_n all lie between the essential minimum and maximum of f.

The above definitions can be generalized slightly. Let μ be a finite signed measure on T. Then we can define its Fourier coefficients by $c_n = \int_{-\pi}^{\pi} e^{-in\theta} d\mu$ and proceed as before with the definitions of T_n and Q_n; in this case

$$Q_n(\mathbf{u}) = \int_{-\pi}^{\pi} |u_0 + u_1 e^{i\theta} + \cdots + u_n e^{in\theta}| \, d\mu.$$

This chapter will be devoted to a proof of the next theorem, which shows that as far as the asymptotic eigenvalue distribution of T_n is concerned, this generalization is very slight indeed.

Theorem 2. *Let μ be a finite signed measure on T, and define the Toeplitz operators T_n as above. Let $\lambda_1^{(n)}, \ldots, \lambda_{n+1}^{(n)}$ denote the eigenvalues of T_n. Let h be the derivative of μ with respect to m. Then if $[a, b] \subset \mathbb{R}$ with $0 \notin [a, b]$ and $m(h^{-1}(\{a\})) = m(h^{-1}(\{b\})) = 0$, we have*

$$\lim_{n\to\infty} \frac{\#\text{ of } \lambda_k^{(n)} \text{ in } [a,b]}{n+1} = m(h^{-1}([a,b]))$$

where m denotes Lebesgue measure on T.

Although stated for closed intervals $[a, b]$, the theorem applies equally well to intervals of the form (a, b), $[a, b)$, $(a, b]$, $(-\infty, a)$, $(-\infty, a]$, (a, ∞), and $[a, \infty)$, provided that they exclude 0 and that any boundary point x satisfies $m(h^{-1}(\{x\})) = 0$.

Several features of this theorem deserve comment. First, it says that the singular part of μ has no effect on the asymptotic eigenvalue distribution of T_n—something that is not at all obvious. Second, since the left-hand side is the fraction of the eigenvalues that are in $[a, b]$ and the right-hand side is the fraction of T that is mapped into $[a, b]$ by h, the theorem can be intuitively understood as saying that (in the limit) the eigenvalues generated by μ spend the same amount of time in any given interval as h does.

We will use the following outline, due to [Grenander and Szegö], [Hoffman], and [Krieger], in proving this theorem. The first three steps will apply to a finite positive measure μ with derivative h with respect to Lebesgue measure (assuming, in steps 1 and 3, that μ is not concentrated on any finite set); the next three will specialize to the case of positive absolutely continuous measures with bounded derivative; finally, we shall show how to extend to any finite absolutely continuous measure, and then to any finite signed measure.

1. The minimum value of Q_n subject to the constraint $u_n = 1$ (or equivalently, $u_0 = 1$) is $\frac{D_n}{D_{n-1}}$ where D_n denotes the determinant of the nth Toeplitz matrix.

2. Szegö's Theorem: If A_0 is the space of continuous $g : T \to \mathbb{C}$ whose negative Fourier coefficients vanish and for which $\int g\, d\theta = 0$, then

$$\inf_{g \in A_0} \int_{-\pi}^{\pi} |1-g|^2\, d\mu = \exp\left\{ \int_{-\pi}^{\pi} \log h\, dm \right\}.$$

3. Since the trigonometric polynomials are dense in A_0, and since the minimum values of Q_n, subject to the constraint mentioned above, form a decreasing sequence, Szegö's theorem implies

$$\lim_{n\to\infty} \frac{D_n}{D_{n-1}} = \exp\left\{ \int_{-\pi}^{\pi} \log h\, dm \right\}.$$

4. We now specialize to the absolutely continuous case, replacing h with f to indicate this new restriction. By an elementary theorem on sequences, if $\frac{D_n}{D_{n-1}}$ converges to a limit, then $(D_n)^{1/(n+1)}$ converges to the same limit. Taking the logarithm of both sides, and using the fact that the determinant of a matrix is the product of its eigenvalues, we have

$$\lim_{n\to\infty} \frac{\log \lambda_1^{(n)} + \cdots + \log \lambda_{n+1}^{(n)}}{n+1} = \int_{-\pi}^{\pi} \log f\, dm.$$

5. Using Vitali's theorem and the fact that the eigenvalues of $I + cT_n$ are $1 + c\lambda_k^{(n)}$, it follows that for any polynomial $p(\lambda)$,

$$\lim_{n\to\infty} \frac{p(\lambda_1^{(n)}) + \cdots + p(\lambda_{n+1}^{(n)})}{n+1} = \int_{-\pi}^{\pi} p(f(\theta))\, dm.$$

   Because of Weierstrass' theorem, the above will then hold if p is allowed to be any continuous function.

6. Since the indicator function of an interval may be approximated well by continuous functions, we can infer the theorem in the case of absolutely continuous measures with bounded derivatives.

7. Using several results from [Krieger], we can extend to arbitrary absolutely continuous measures, and then show that the addition of a singular part does not affect the limiting distribution.

We now give details for each part in turn.

2.1 Minimum Value of Q_n

2.1.1 Orthogonal Polynomials on the Unit Circle

Using the formula for Gram-Schmidt orthogonalization (see Appendix B) , we can find a family of polynomials $\phi_n(z)$ that are orthogonal with respect to the inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(e^{i\theta}) \overline{g(e^{i\theta})}\, d\mu.$$

To do this, we just use the facts that $1, z, \ldots, z^n$ form a basis for polynomials of degree n or less, and that $\langle z^l, z^m \rangle = \int_{-\pi}^{\pi} e^{il\theta} e^{-im\theta}\, d\mu = c_{l-m}$. (We assume here that the monomials are independent in $L^p(\mu)$. This fails to be the case iff μ is concentrated on a finite set, in which case $L^p(\mu)$ is finite-dimensional; otherwise, any monomial can only be equal to a lower-degree polynomial on a finite set, which is not almost everywhere. For example, if μ has any absolutely continuous part we need not worry.) Thus, an orthonormal basis for the same space is given by

$$\begin{aligned}
\phi_n(z) &= \frac{1}{\sqrt{D_{n-1} D_n}} \begin{vmatrix} c_0 & c_{-1} & c_{-2} & \ldots & c_{-n} \\ c_1 & c_0 & c_{-1} & \ldots & c_{-n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & c_{n-3} & \ldots & c_{-1} \\ 1 & z & z^2 & \ldots & z^n \end{vmatrix} \qquad n \geq 1 \\
\phi_0(z) &= \frac{1}{\sqrt{c_0}}
\end{aligned}$$

where

$$D_n = \begin{vmatrix} c_0 & c_{-1} & c_{-2} & \cdots & c_{-n} \\ c_1 & c_0 & c_{-1} & \cdots & c_{-n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & c_{n-3} & \cdots & c_{-1} \\ c_n & c_{n-1} & c_{n-2} & \cdots & c_0 \end{vmatrix} \qquad n \geq 0.$$

Note that the leading coefficient of ϕ_n is $\sqrt{\frac{D_{n-1}}{D_n}}$ for $n \geq 1$.

2.1.2 Minimizing the Toeplitz Form

We are now in a position to find the minimum value of

$$Q_n(\mathbf{u}) = \int_{-\pi}^{\pi} |u_0 + u_1 e^{i\theta} + \cdots + u_n e^{in\theta}|^2 \, d\mu$$

subject to the constraint $u_n = 1$. Let $g(z) = z^n + u_{n-1} z^{n-1} + \cdots + u_1 z + u_0$. Because the ϕ's form a basis, we can write

$$g(z) = a_0 \phi_0(z) + \cdots + a_n \phi_n(z).$$

Because ϕ_n is the only one of $\phi_0, \ldots, \phi_n$ that has a z^n term, and since the z^n coefficient was previously found to be $\sqrt{\frac{D_{n-1}}{D_n}}$, we can see that $a_n = \sqrt{\frac{D_n}{D_{n-1}}}$. Now by orthogonality we have

$$\int_{-\pi}^{\pi} |u_0 + u_1 e^{i\theta} + \cdots + u_n e^{in\theta}|^2 \, d\mu = \langle g, g \rangle = |a_0|^2 + \cdots + |a_n|^2 \geq |a_n|^2 = \frac{D_n}{D_{n-1}}.$$

This shows that for $u_n = 1$, $Q_n(\mathbf{u}) \geq \frac{D_n}{D_{n-1}}$; moreover, this minimum can actually be attained by setting $g(z) = z^n + u_{n-1} z^{n-1} + \cdots + u_1 z + u_0 = \sqrt{\frac{D_n}{D_{n-1}}} \phi_n(z)$.

It is easy to see that the same minimum applies if the constraint is $u_0 = 1$ instead, since

$$\begin{aligned} |u_n e^{in\theta} + \cdots + u_1 e^{i\theta} + u_0| &= |\overline{u_n} e^{-in\theta} + \cdots + \overline{u_1} e^{-i\theta} + \overline{u_0}| \\ &= |\overline{u_n} + \cdots + \overline{u_1} e^{i(n-1)\theta} + \overline{u_0} e^{in\theta}|. \end{aligned}$$

It is the latter constraint which we shall make use of later.

Side Note for the Curious

We could also have minimized Q_n subject to the constraint $u_0 = 1$ using the following technique: Let $d_0, \ldots, d_n$ be the constant terms of $\phi_0, \ldots, \phi_n$. Then the constraint $u_0 = 1$ becomes $a_0 d_0 + \cdots + a_n d_n = 1$, or $\mathbf{a} \cdot \bar{\mathbf{d}} = 1$ where we view $\mathbf{a} = (a_0, \ldots, a_n)$ and $\mathbf{d} = (d_0, \ldots, d_n)$ as vectors in $\mathbb{C}^{n+1}$. We are then trying to minimize $\|\mathbf{a}\|$ with this constraint. Clearly this is accomplished by choosing $\mathbf{a}$ to be in the same direction as $\bar{\mathbf{d}}$, i.e. $\mathbf{a} = \frac{\bar{\mathbf{d}}}{\mathbf{d} \cdot \mathbf{d}}$ and $\langle g, g \rangle = \|\mathbf{a}\|^2 = \frac{1}{\mathbf{d} \cdot \mathbf{d}}$.

Since we have found two ways of minimizing the quadratic form, we have the following theorem:

Theorem 3. *Let $\{c_n\}_{n=-\infty}^{\infty}$ be any list of complex numbers such that $\overline{c_n} = c_{-n}$. Then if D_n denotes the determinant of the nth Toeplitz matrix generated by the c_n (i.e. the matrix T with $T_{i,j} = c_{i-j}$, $i, j = 0, \ldots, n$), and $\tilde{D_n}$ denotes $\frac{1}{\sqrt{D_n D_{n-1}}}$ times the determinant of the upper right minor of the nth Toeplitz matrix (i.e. the one formed by crossing out the left column and the bottom row), then*

$$\frac{D_n}{D_{n-1}} = \frac{1}{|\tilde{D_0}|^2 + \cdots + |\tilde{D_n}|^2}.$$

This follows from simply observing that $\tilde{D_k}$ is the constant term in ϕ_k in the discussion above. For example, this proves that for any $c_0, c_1, c_2, c_3 \in \mathbb{C}$,

$$\frac{\begin{vmatrix} c_0 & \bar{c_1} & \bar{c_2} \\ c_1 & c_0 & \bar{c_1} \\ c_2 & c_1 & c_0 \end{vmatrix}}{\begin{vmatrix} c_0 & \bar{c_1} & \bar{c_2} & \bar{c_3} \\ c_1 & c_0 & \bar{c_1} & \bar{c_2} \\ c_2 & c_1 & c_0 & \bar{c_1} \\ c_3 & c_2 & c_1 & c_0 \end{vmatrix}} = \frac{1}{c_0} \left|1\right|^2 + \frac{1}{c_0 \begin{vmatrix} c_0 & \bar{c_1} \\ c_1 & c_0 \end{vmatrix}} \left|\bar{c_1}\right|^2 + \frac{1}{\begin{vmatrix} c_0 & \bar{c_1} \\ c_1 & c_0 \end{vmatrix} \begin{vmatrix} c_0 & \bar{c_1} & \bar{c_2} \\ c_1 & c_0 & \bar{c_1} \\ c_2 & c_1 & c_0 \end{vmatrix}} \begin{vmatrix} \bar{c_1} & \bar{c_2} \\ c_0 & \bar{c_1} \end{vmatrix}^2$$
$$+ \frac{1}{\begin{vmatrix} c_0 & \bar{c_1} & \bar{c_2} & \bar{c_3} \\ c_1 & c_0 & \bar{c_1} & \bar{c_2} \\ c_2 & c_1 & c_0 & \bar{c_1} \\ c_3 & c_2 & c_1 & c_0 \end{vmatrix} \begin{vmatrix} c_0 & \bar{c_1} & \bar{c_2} \\ c_1 & c_0 & \bar{c_1} \\ c_2 & c_1 & c_0 \end{vmatrix}} \begin{vmatrix} \bar{c_1} & \bar{c_2} & \bar{c_3} \\ c_0 & \bar{c_1} & \bar{c_2} \\ c_1 & c_0 & \bar{c_1} \end{vmatrix}^2$$

where it is understood that the squares of determinants actually refer to the squares of their absolute values.

The reader is challenged to prove this directly!

2.2 Szegö's Theorem

Definition 2. *Let A denote the set of continuous functions $f : T \to \mathbb{C}$ such that*

$$\int_{-\pi}^{\pi} e^{in\theta} f \, dm = 0 \qquad n = 1, 2, 3, \ldots$$

and $A_0 = \{f \in A \mid \int_{-\pi}^{\pi} f \, dm = 0\}$.

Each $f \in A$ defines an analytic function $\tilde{f}$ inside the unit disk, by the Poisson integral:

$$\tilde{f}(re^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{it}) \frac{1 - r^2}{1 - 2r\cos\theta + r^2} \, dt.$$

See [Duren] for details. Note that $\tilde{f}(0) = \int f \, dm$; in particular, A_0 corresponds to those functions which are analytic inside the unit disc and vanish at the origin. This correspondence between f and $\tilde{f}$ allows us to see, for example, that A and A_0 are both algebras over $\mathbb{C}$.

Theorem 4. *The set TP of trigonometric polynomials of the form*

$$P(\theta) = \sum_{k=0}^{n} a_k e^{ik\theta}$$

is a uniformly dense subset of A.

Proof. It is clear that $TP \subset A$. Now every $f \in A$, being continuous, is the uniform limit of the Cesàro means of its Fourier series; but these Cesàro means are in TP, so TP is dense in A. □

Remark. We can use exactly the same reasoning to see that the set $TP_0 \subset TP$, consisting of all trigonometric polynomials in TP for which the constant term vanishes, is uniformly dense in A_0.

Theorem 5. *The real parts of functions in A are uniformly dense in the real-valued continuous functions on T.*

Proof. A includes all trigonometric polynomials of the form

$$P(\theta) = \sum_{k=0}^{n} a_k e^{ik\theta},$$

so Re A include all trigonometric polynomials of the form

$$Q(\theta) = \sum_{k=-n}^{n} c_k e^{ik\theta}, \qquad c_{-k} = \bar{c}_k$$

that is, all real-valued trigonometric polynomials. Such polynomials are dense in the real-valued continuous functions on T, since any continuous $f : T \to \mathbb{R}$ is the uniform limit of the Cesàro means of its Fourier series, and the Cesàro means are real-valued trigonometric polynomials. (Alternatively, the real-valued trigonometric polynomials are easily shown to be a separating subalgebra of $C(T)$ which vanishes nowhere, so they are dense by the Stone-Weierstrass theorem.) □

Corollary 1. *If μ is a finite real measure on T such that $\int f d\mu = 0$ for all $f \in A_0$, then μ is a constant multiple of Lebesgue measure.*

Proof. Let $d\mu_1 = d\mu - \lambda\, dm$ where $\lambda = \int d\mu$. Then μ_1 is a finite real measure with the property that $\int f\, d\mu_1 = \int f\, d\mu - \lambda \int f\, dm = 0 - 0 = 0$ for all $f \in A_0$. Thus μ_1 is "orthogonal" to every $f \in A_0$; by Theorem 5, μ_1 must be orthogonal to every continuous $f : T \to \mathbb{R}$ and must therefore be the zero measure. Hence $d\mu = \lambda\, dm$. □

Theorem 6. *Let μ be a finite positive measure on T and let F be the projection of the constant function 1 onto the closed subspace of $L^2(\mu)$ spanned by A_0. Let μ_s be the singular part of μ. Then $1 - F$ vanishes almost everywhere with respect to μ_s.*

Proof. Let S be the closed subspace of $L^2(\mu)$ spanned by A_0. Then $1 - F$ is orthogonal to S. Let f_n be a sequence of functions in A_0 converging to F. Now for any fixed $f \in A_0$, $f(1 - f_n)$ is in $A_0 \subset S$ and therefore orthogonal to $1 - F$. By the continuity of inner products, this implies that $f(1 - F)$ is also orthogonal to $1 - F$, i.e.

$$\int f |1 - F|^2\, d\mu = 0, \qquad f \in A_0.$$

By Corollary 1, this implies that $|1 - F|^2 d\mu$ is a constant multiple of Lebesgue measure; in particular, $|1 - F|^2 d\mu$ is absolutely continuous.

Now suppose $\mu = \mu_a + \mu_s$ where μ_a is absolutely continuous and μ_s is singular. Let B be a set on which μ_s is concentrated and for which $m(B) = 0$. Let $A \subset B$. Then $\int_A |1 - F|^2\, d\mu = 0$ by the absolute continuity of $|1 - F|^2 d\mu$. But we also have $\int_A |1 - F|^2\, d\mu = \int_A |1 - F|^2\, d\mu_a + \int_A |1 - F|^2\, d\mu_s = \int_A |1 - F|^2\, d\mu_s$. Hence $\int_A |1 - F|^2\, d\mu_s = 0$ for any subset $A \subset B$. This implies that $1 - F$ vanishes almost everywhere with respect to μ_s. □

Corollary 2. *If μ is a positive measure on T with absolutely continuous part μ_a, then*

$$\inf_{f \in A_0} \int |1-f|^2 \, d\mu = \inf_{f \in A_0} \int |1-f|^2 \, d\mu_a.$$

Proof. As above, we let F be the orthogonal projection of 1 onto the closed subspace of $L^2(\mu)$ spanned by A_0. Then the infimum on the left above is just $\int |1-F|^2 d\mu$. Now for any $f \in A_0$,

$$\int f(1-F)\, d\mu_a = \int f(1-F)\, d\mu - \int f(1-F)\, d\mu_s = \int f(1-F)\, d\mu = 0$$

so $1-F$ is also orthogonal to A_0 in $L^2(\mu_a)$. Hence F is also the orthogonal projection of 1 onto the closed span of A_0 in $L^2(\mu_a)$, which implies that

$$\begin{aligned}\inf_{f \in A_0} \int |1-f|^2 \, d\mu_a = \int |1-F|^2 d\mu_a &= \int |1-F|^2 d\mu - \int |1-F|^2 d\mu_s \\ &= \int |1-F|^2 d\mu = \inf_{f \in A_0} \int |1-f|^2 \, d\mu.\end{aligned}$$

□

Theorem 7. *Let $h : T \to \mathbb{R}$ be a nonnegative integrable function. Then*

$$\inf_{f \in A_0} \int h e^{\operatorname{Re} f} \, dm = \exp\left\{ \int \log h \, dm \right\}.$$

The right-hand side is to be interpreted as 0 *if* $\log h$ *is not integrable.*

Proof. First we treat the case where $\log h$ is integrable. Let I_0 denote the set of integrable $g : T \to \mathbb{R}$ with $\int g \, dm = 0$. Note that $\{\operatorname{Re} f \mid f \in A_0\} \subset I_0$ since for $f \in A_0$, $\int \operatorname{Re} f \, dm = \operatorname{Re} \int f \, dm = 0$. Now for any $g \in I_0$ we have by convexity

$$\begin{aligned}\exp\left\{ \int \log h \, dm \right\} &= \exp\left\{ \int (\log h + g)\, dm \right\} \\ &\leq \int \exp\{\log h + g\}\, dm = \int h e^g dm.\end{aligned}$$

Thus,

$$\exp\left\{ \int \log h \, dm \right\} \leq \inf_{g \in I_0} \int h e^g dm.$$

If we let $\lambda = \int \log h \, dm$ and $G = \lambda - \log h$, then $G \in I_0$ and

$$\int he^G dm = \int e^\lambda dm = \exp\left\{\int \log h \, dm\right\}.$$

Thus, we have

$$\exp\left\{\int \log h \, dm\right\} = \inf_{g \in I_0} \int he^g dm$$

and this infimum is actually attained at G.

Next we show that

$$\inf_{g \in I_0} \int he^g dm = \inf_{f \in A_0} \int he^{\mathrm{Re}\, f} dm.$$

Because $\{\mathrm{Re}\, f \mid f \in A_0\} \subset I_0$, we clearly have

$$\inf_{g \in I_0} \int he^g d\theta \leq \inf_{f \in A_0} \int he^{\mathrm{Re}\, f} d\theta.$$

Now let $g \in I_0$, let g^+ and g^- be its positive and negative parts, and define a sequence of functions g_n as follows:

$$g_n^+(x) = \begin{cases} g^+(x) & \text{if } g^+(x) \leq n \\ n & \text{if } g^+(x) \geq n \end{cases}$$

$$g_n^-(x) = \begin{cases} g^-(x) & \text{if } g^-(x) \leq k_n \\ k_n & \text{if } g^-(x) \geq k_n \end{cases}$$

where k_n is chosen such that $\int g_n \, dm = 0$. (We can choose such a k_n because on the compact set T, the integral of g_n^- is a uniformly continuous function of the truncation parameter k_n which is 0 for $k_n = 0$ and approaches $\int g^- \, dm = \int g^+ \, dm$ as $k_n \to \infty$; by the intermediate value theorem we can choose a k_n for which this integral attains the intermediate value $\int g_n^+ \, dm$.) Then g_n is bounded, $g_n^+ \nearrow g^+$, and $g_n^- \nearrow g^-$. By Lebesgue's Monotone Convergence Theorem, $\int he^{g_n^+} \, dm \to \int he^{g^+} \, dm$. By Lebesgue's Dominated Convergence Theorem, $\int he^{-g_n^-} \, dm \to \int he^{-g^-} \, dm$. Hence $\int he^{g_n} \, dm \to \int he^g \, dm$ and

$$\inf_{g \in I_0} \int he^g dm = \inf_{\substack{g \in I_0 \\ g \text{ bounded}}} \int he^g dm.$$

Now any integrable g is the pointwise almost everywhere limit of a sequence of continuous functions (since continuous functions are dense in L^1, and L^1 convergence implies convergence in measure which implies a.e. convergence of a subsequence); for g bounded, since $\operatorname{Re} A$ is dense in the continuous functions, g is the pointwise almost everywhere limit of a bounded sequence of functions $f_n \in \operatorname{Re} A$. By Lebesgue's Dominated Convergence Theorem, $\int he^{f_n} \to \int he^g$ so

$$\inf_{f \in A_0} \int he^{\operatorname{Re} f} dm = \inf_{\substack{g \in I_0 \\ g \text{ bounded}}} \int he^g dm = \inf_{g \in I_0} \int he^g dm$$

This proves the theorem in the case where $\log h$ is integrable. If $\log h$ is not integrable, we must have $\int \log h \, dm = -\infty$ since h is integrable and $\log h \leq h$. Now for any $\epsilon > 0$, $\log(h + \epsilon)$ is integrable because it's bounded below by $\log \epsilon$ and above by $\log(h(1+\epsilon)) = \log h + \log(1+\epsilon)$ for $h \geq 1$ and $\log(1+\epsilon)$ for $h \leq 1$. Then

$$\exp\left\{\int \log(h+\epsilon)\, dm\right\} = \inf_{f \in A_0} \int (h+\epsilon) e^{\operatorname{Re} f} dm.$$

Consider what happens as $\epsilon \to 0$. Since the negative parts of $\log(h+\epsilon)$ increase monotonically to the negative part of $\log(h)$ and the integrals of the positive parts are decreasing, $\int \log(h+\epsilon)\, dm \to -\infty$ as $\epsilon \to 0$, so the left-hand side approaches 0. The right-hand side is at least

$$\inf_{f \in A_0} \int he^{\operatorname{Re} f} dm + \epsilon \inf_{f \in A_0} \int e^{\operatorname{Re} f} dm$$

which clearly approaches

$$\inf_{f \in A_0} \int he^{\operatorname{Re} f} dm$$

as $\epsilon \to 0$. Thus, the theorem holds for nonintegrable $\log h$ as well. $\square$

Theorem 8 (Szegö's Theorem)**.** *Let μ be a finite positive measure on T and let h be the derivative of μ with respect to normalized Lebesgue measure m. Then*

$$\inf_{f \in A_0} \int_{-\pi}^{\pi} |1 - f|^2 \, d\mu = \exp\left\{\int_{-\pi}^{\pi} \log h \, dm\right\}.$$

Proof. Since $g \in A_0 \Leftrightarrow 2g \in A_0$, Theorem 7 implies

$$\exp\left\{\int \log h \, dm\right\} = \inf_{g \in A_0} \int he^{2\operatorname{Re} g} \, dm.$$

For $g \in A_0$, e^g defines, by the Poisson integral, an analytic function whose value at the origin is 1 (since g defines an analytic function $\tilde{g}$ with $\tilde{g}(0) = 0$). Thus, $e^g = 1 - f$ for some $f \in A_0$. (See Appendix C for details). Then $e^{2\operatorname{Re} g} = |1-f|^2$ for some $f \in A_0$. Thus

$$\exp\left\{\int \log h \, dm\right\} \geq \inf_{f \in A_0} \int h|1-f|^2 \, dm.$$

We obtain the reverse inequality by applying this one to a different h. Let $g \in A_0$ and replace h with $|1-g|^2$; then the last inequality becomes

$$\exp\left\{\int \log|1-g|^2 dm\right\} \geq \inf_{f \in A_0} \int |1-f-g+fg|^2 \, dm \geq 1$$

since $f + g - fg \in A_0$ and hence is orthogonal to 1 in $L^2(m)$. Thus, $\log|1-g|^2$ is integrable and $\int \log|1-g|^2 \geq 0$. This allows us to write $|1-g|^2 = ke^p$ where p is a real L^1 function with $\int p \, dm = 0$ and the constant $k \geq 1$. Going back to our original h, we see that

$$\int |1-g|^2 h \, d\theta = k \int he^p \, dm \geq \inf_{p \in I_0} \frac{1}{2\pi} \int he^p \, dm = \exp\left\{\int \log h \, dm\right\}.$$

Taking the infimum over g, we have

$$\inf_{g \in A_0} \int |1-g|^2 h \, dm \geq \exp\left\{\int \log h \, dm\right\}.$$

Putting together the two inequalities, we have

$$\inf_{g \in A_0} \int |1-g|^2 h \, dm = \exp\left\{\int \log h \, dm\right\}.$$

By Corollary 2, this infimum is equal to

$$\inf_{g \in A_0} \int |1-g|^2 \, d\mu$$

which completes the proof. □

2.3 Connecting Szegö to Toeplitz

By the remark following theorem 4, any $g \in A_0$ is the uniform limit of functions in the subspace

$$TP_0 = \left\{\sum_{k=1}^{n} a_k e^{ik\theta} \mid a_k \in \mathbb{C}\right\}$$

so that we may replace the conclusion of Szegö's theorem by

$$\inf_{g \in TP_0} \int |1-g|^2 \, d\mu = \exp \left\{ \int_{-\pi}^{\pi} \log h \, dm \right\}.$$

Now $\{1 - g \mid g \in TP_0\}$ is just the set of trigonometric polynomials with constant term equal to 1; we showed in section 2.1 that

$$\inf_{\substack{p \in TP \\ \text{constant term of } p = 1 \\ \text{degree}(p) \leq n}} \int |p|^2 \, d\mu = \frac{D_n}{D_{n-1}}.$$

Since the infimum is taken over a strictly larger set as n increases, $\frac{D_n}{D_{n-1}}$ is monotonically decreasing and

$$\lim_{n \to \infty} \frac{D_n}{D_{n-1}} = \inf_{\substack{h \in TP \\ \text{constant term of } h = 1}} \int |h|^2 \, d\mu$$

$$= \inf_{g \in TP_0} \int |1-g|^2 \, d\mu = \exp \left\{ \int_{-\pi}^{\pi} \log h \, dm \right\}.$$

2.4 Ratios to Roots

For any nonnegative real sequence b_n, if $\frac{b_{n+1}}{b_n}$ converges to some limit L, then $(b_n)^{1/n}$ also converges to L. This is a standard theorem from calculus [Rudin (1976)]. One can easily see that $(b_n)^{1/(n+1)}$ converges to L as well. Applying this to the sequence D_n of Toeplitz determinants, we see that

$$\lim_{n \to \infty} (D_n)^{1/(n+1)} = \exp \left\{ \int_{-\pi}^{\pi} \log h \, dm \right\}.$$

Suppose now that the measure generating our Toeplitz matrices is such that all the eigenvalues of all the matrices are strictly positive; then we can take the logarithm of both sides, yielding

$$\lim_{n \to \infty} \frac{\log \lambda_1^{(n)} + \cdots + \log \lambda_{n+1}^{(n)}}{n+1} = \int_{-\pi}^{\pi} \log h \, dm.$$

For example, suppose f is essentially bounded, say $|f| \leq M$ a.e. Then for $|z| \leq \frac{1}{M}$, $1 + zf$ is positive a.e., so the eigenvalues of the Toeplitz form generated by $1+zf$, which are just $1+z\lambda_1^{(n)},\ldots,1+z\lambda_{n+1}^{(n)}$, are positive. Thus we have

$$\lim_{n\to\infty}\frac{\log(1+z\lambda_1^{(n)})+\cdots+\log(1+z\lambda_{n+1}^{(n)})}{n+1}=\int_{-\pi}^{\pi}\log(1+zf)\,dm$$

for $|z|\le\frac{1}{M}$.

2.5 Vitali and Weierstrass

We can restate the conclusion of the previous section in terms of the following definition.

Definition 3. *Let $f:T\to\mathbb{R}$ be an essentially bounded function with $|f|\le M$ a.e., and late $\lambda_k^{(n)}$ denote the eigenvalues of the Toeplitz forms generated by f. We define $\mathfrak{S}$ to be the set of functions $F:[-M,M]\to\mathbb{R}$ such that*

$$\lim_{n\to\infty}\frac{F(\lambda_1^{(n)})+\cdots+F(\lambda_{n+1}^{(n)})}{n+1}=\int_{-\pi}^{\pi}F(f(\theta))\,dm.$$

Then we have just finished proving that for all $z\in\mathbb{R}$ with $|z|\le\frac{1}{M}$, $F_z(x)=\log(1+zf)\in\mathfrak{S}$. Our next objective is to show that this actually implies that $\mathfrak{S}$ contains all continuous functions on $[-M,M]$.

Let

$$g_n(z)=\frac{\log(1+z\lambda_1^{(n)})+\cdots+\log(1+z\lambda_{n+1}^{(n)})}{n+1}-\int_{-\pi}^{\pi}\log(1+zf)\,dm.$$

Then g_n is analytic in the open disk $|z|<\frac{1}{M}$ in the complex plane (since none of the arguments of the logarithms can be negative real). Also note that g_n is uniformly bounded in n and z on any subdisk; for example, on $|z|\le\frac{1}{2M}$ we have $|\log(1+z\lambda_k^{(n)})|<\log(3/2)$ and $|\log(1+zf)|<\log(3/2)$ so $|g_n(z)|\le 2\log(3/2)$ on this subdisk. By Vitali's theorem, since g_n tends to a limit on the intersection of this subdisk with the real line, it tends uniformly to an analytic function $g(z)$ on any sub-subdisk. Since g is analytic and equals zero on the intersection of the real line with this sub-subdisk, it must be identically zero. Thus, $g_n\to 0$ uniformly on a disk containing the origin, which implies that any given Taylor coefficient of f_n will approach 0 as $n\to\infty$. But for $k\ge 1$, the kth Taylor coefficient of g_n is

$$\frac{(-1)^{k+1}}{k}\left(\frac{(\lambda_1^{(n)})^k+\cdots+(\lambda_{n+1}^{(n)})^k}{n+1}-\int_{-\pi}^{\pi}(f(\theta))^k\,dm\right)$$

which approaches zero as $n \to \infty$ iff

$$\lim_{n\to\infty} \frac{(\lambda_1^{(n)})^k + \cdots + (\lambda_{n+1}^{(n)})^k}{n+1} = \int_{-\pi}^{\pi} (f(\theta))^k \, dm.$$

This shows that $F(x) = x^k$ is in $\mathfrak{S}$ for $k \geq 1$. Since constant functions are obviously in $\mathfrak{S}$ as well, we see that $\mathfrak{S}$ contains all monomials. It is also easy to see that $\mathfrak{S}$ is a vector space over $\mathbb{R}$, so that it must contain all polynomials. By Weierstrass' theorem, in order to show that it contains all continuous functions, we need only show that it is uniformly closed.

Let $\epsilon > 0$. Suppose F is in the closure of $\mathfrak{S}$, and let $g \in \mathfrak{S}$ such that $|F - g| \leq \frac{\epsilon}{3}$ uniformly. Choose N such that

$$\left| \frac{g(\lambda_1^{(n)}) + \cdots + g(\lambda_{n+1}^{(n)})}{n+1} - \int_{-\pi}^{\pi} g(f(\theta)) \, dm \right| < \frac{\epsilon}{3} \qquad \text{for } n > N.$$

Then for $n > N$,

$$\begin{aligned}
&\left| \frac{\sum_{k=1}^{n+1} F(\lambda_k^{(n)})}{n+1} - \int_{-\pi}^{\pi} F(f) \, dm \right| \\
&= \left| \frac{\sum_{k=1}^{n+1} g(\lambda_k^{(n)})}{n+1} - \int_{-\pi}^{\pi} g(f) \, dm + \frac{\sum_{k=1}^{n+1} F(\lambda_k^{(n)}) - g(\lambda_k^{(n)})}{n+1} + \int_{-\pi}^{\pi} (g(f) - F(f)) \, dm \right| \\
&\leq \left| \frac{\sum_{k=1}^{n+1} g(\lambda_k^{(n)})}{n+1} - \int_{-\pi}^{\pi} g(f) \, dm \right| + \frac{\sum_{k=1}^{n+1} |F(\lambda_k^{(n)}) - g(\lambda_k^{(n)})|}{n+1} + \int_{-\pi}^{\pi} |g(f) - F(f)| \, dm \\
&\qquad < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

Thus,

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n+1} F(\lambda_k^{(n)})}{n+1} = \int_{-\pi}^{\pi} F(f(\theta)) \, dm.$$

That is to say, $F \in \mathfrak{S}$. This completes the proof that $\mathfrak{S}$ is uniformly closed and therefore contains all continuous real functions on $[-M, M]$. $\square$

2.6 From Continuous Functions to Indicators

We will now demonstrate that indicator functions of intervals are in $\mathfrak{S}$ as well. Let $I = [a, b]$ be some interval for which $m(f^{-1}(\{a\})) = m(f^{-1}(\{b\})) =$

0. (Our analysis will apply equally well to other types of intervals for which the boundary points satisfy this condition; we use a finite closed interval for simplicity.) Let $\chi_I(x)$ be the indicator function for this interval. We approximate χ_I by continuous functions as follows:

$$g_m(x) = \begin{cases} 0 & \text{if } x \leq a - \frac{1}{m} \\ m(x-a)+1 & \text{if } a - \frac{1}{m} < x \leq a \\ 1 & \text{if } a < x \leq b \\ m(b-x)+1 & \text{if } b < x \leq b + \frac{1}{m} \\ 0 & \text{if } x > b + \frac{1}{m} \end{cases}$$

$$\tilde{g}_m(x) = \begin{cases} 0 & \text{if } x \leq a \\ m(x-a) & \text{if } a < x \leq a + \frac{1}{m} \\ 1 & \text{if } a + \frac{1}{m} < x \leq b - \frac{1}{m} \\ m(b-x) & \text{if } b - \frac{1}{m} < x \leq b \\ 0 & \text{if } x > b \end{cases}$$

Note that, depending on the values of a and b, these may only be well-defined for sufficiently large m; this of course presents no problem. These approximating functions satisfy

- $0 \leq \tilde{g}_k \leq \chi_I \leq g_l \leq 1 \qquad \forall k, l \in \mathbb{N}$
- $g_m = \chi_I$ except on $(a - \frac{1}{m}, a) \cup (b, b + \frac{1}{m})$
- $\tilde{g}_m = \chi_I$ except on $[a, a + \frac{1}{m}) \cup (b - \frac{1}{m}, b]$

Now let $\epsilon > 0$. Choose some open set $U \supset \{a, b\}$ with $m(f^{-1}(U)) < \frac{\epsilon}{2}$. (To be perfectly clear, U contains the points a and b separately but not generally the interval between them.) Then $\exists M \in \mathbb{N}$ such that $\forall m \geq M$, $g_m = \chi_I = \tilde{g}_m$ outside U. We have

$$\int_{-\pi}^{\pi} g_M(f(\theta))\, dm = \int_{-\pi}^{\pi} \chi_I(f(\theta))\, dm + \int_{-\pi}^{\pi} \left(g_M(f(\theta)) - \chi_I(f(\theta))\right)\, dm.$$

Now $0 \leq g_M(f(\theta)) - \chi_I(f(\theta)) \leq 1$ and is equal to 0 outside $f^{-1}(U)$, so

$$0 \leq \int_{-\pi}^{\pi} \left(g_M(f(\theta)) - \chi_I(f(\theta))\right)\, dm \leq \epsilon.$$

Similarly, we have

$$-\epsilon \leq \int_{-\pi}^{\pi} \left(\tilde{g}_M(f(\theta)) - \chi_I(f(\theta))\right)\, dm \leq 0.$$

Now since $\tilde{g}_M \leq \chi_I \leq g_M$ everywhere,

$$\frac{\sum_{k=1}^{n+1} \tilde{g}_M(\lambda_k^{(n)})}{n+1} \leq \frac{\sum_{k=1}^{n+1} \chi_I(\lambda_k^{(n)})}{n+1} \leq \frac{\sum_{k=1}^{n+1} g_M(\lambda_k^{(n)})}{n+1}$$

for all n.

By section 2.5, we have, for sufficiently large n,

$$\frac{\sum_{k=1}^{n+1} \tilde{g}_M(\lambda_k^{(n)})}{n+1} > \int_{-\pi}^{\pi} \tilde{g}_M(f)\, dm - \frac{\epsilon}{2}$$

and

$$\frac{\sum_{k=1}^{n+1} g_M(\lambda_k^{(n)})}{n+1} < \int_{-\pi}^{\pi} g_M(f)\, dm + \frac{\epsilon}{2}$$

so that

$$\begin{aligned} \int_{-\pi}^{\pi} \chi_I(f)\, dm - \epsilon \leq \int_{-\pi}^{\pi} \tilde{g}_M(f)\, dm - \frac{\epsilon}{2} < \frac{\sum_{k=1}^{n+1} \tilde{g}_M(\lambda_k^{(n)})}{n+1} \\ \leq \frac{\sum_{k=1}^{n+1} \chi_I(\lambda_k^{(n)})}{n+1} \leq \frac{\sum_{k=1}^{n+1} g_M(\lambda_k^{(n)})}{n+1} < \int_{-\pi}^{\pi} g_M(f)\, dm + \frac{\epsilon}{2} \leq \int_{-\pi}^{\pi} \chi_I(f)\, dm + \epsilon \\ \implies \left| \frac{\sum_{k=1}^{n+1} \chi_I(\lambda_k^{(n)})}{n+1} - \int_{-\pi}^{\pi} \chi_I(f)\, dm \right| < \epsilon. \end{aligned}$$

Thus, indicator functions of intervals are also in $\mathfrak{S}$, as long as the endpoints have negligible preimages. This proves our main theorem (2) for absolutely continuous positive measures with bounded derivative.

2.7 Extending to Finite Measures

So far we have proved the main theorem in the case of Toeplitz operators generated by bounded measurable functions. Theorem 6 in [Krieger] shows that this may be extended to real $f \in L^1(T)$. (In fact, it is true for a larger class of functions, which includes every $L^p(T)$ for $1 \leq p < \infty$.) In terms of measures, this establishes it for absolutely continuous positive finite measures. We now show that the addition of a singular part will not affect the limiting distribution.

First, suppose ν is a finite positive singular measure and $z > 0$ any positive real, and consider $\eta = m + z\nu$. It is easy to check that the nth Toeplitz matrix generated by η is the identity plus the matrix generated by

$z\nu$, and hence has eigenvalues $1 + z\lambda_j^{(n)}$. Now these are all positive (in fact, at least 1), so we can apply the result of section 2.4 to conclude that

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=1}^{n+1} \log(1 + z\lambda_k^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log 1 d\theta = 0.$$

Now fix any $a > 0$. For $z \geq \frac{e-1}{a}$, $\log(1 + zt) \geq \chi_{[a,\infty)}(t) \geq 0$ for $t \in (0, \infty)$ so that we have

$$\lim_{n\to\infty} \frac{\text{\# of } \lambda_l^{(n)} \text{ in } [a, \infty)}{n+1} = \lim_{n\to\infty} \frac{1}{n+1} \sum_{k=1}^{n+1} \chi_{[a,\infty)}(\lambda_k^{(n)}) = 0.$$

Since any interval excluding zero is contained in a half-infinite interval excluding zero,

$$\lim_{n\to\infty} \frac{\text{\# of } \lambda_l^{(n)} \text{ in } I}{n+1} = 0$$

for any interval I with $0 \notin I$. This means that the eigenvalue distribution of a positive singular measure "goes to zero" in an appropriate sense; even though a few eigenvalues may remain large or even tend to infinity, relatively fewer and fewer can do so.

In order to apply this, we take advantage of the following theorem from [Krieger]:

Theorem 9. *Let T_n, T_n^1, T_n^2 be sequences of $(n+1)$ by $(n+1)$ Hermitian matrices with $T_n = T_1 + T_2$. Define*

$$D_n(t) = \begin{cases} \frac{1}{n+1}\{\text{\# eigenvalues of } T_n \text{ less than } t\} \text{ for } t < 0 \\ -\frac{1}{n+1}\{\text{\# eigenvalues of } T_n \text{ greater than } t\} \text{ for } t > 0 \end{cases}$$

and similarly for $D_n^1(t)$ and $D_n^2(t)$. If there is a function D increasing on $(-\infty, 0)$ and $(0, \infty)$ with $D_n^1(t) \to D(t)$ at every continuity point of D, and if $D_n^2(t) \to 0$ for all $t \neq 0$, then $D_n(t) \to D(t)$ at every continuity point of D.

We apply it twice: First, a finite signed singular measure is the difference of positive singular measures, by the Jordan decomposition; this theorem tells us that an arbitrary singular measure also has an eigenvalue distribution that tends to zero in the above sense. (Explicitly, let $D(t) = 0$ everywhere; for any interval I excluding zero, choose a half-interval $(-\infty, t)$ or (t, ∞) containing I, and apply the theorem to find that

$$\lim_{n\to\infty} \frac{\text{\# of eigenvalues generated by } \mu_s \text{ in } I}{n+1} = 0$$

for the singular measure μ_s.) Finally, let μ be any finite signed measure on T. Then $\mu = \mu_a + \mu_s$ where μ_a is absolutely continuous and μ_s singular. Since the eigenvalue distribution generated by μ_s tends to zero, and the distribution generated by μ_a follows the main theorem (the derivative being in $L^1(T)$), we have finished proving the main theorem.

Chapter 3

The q-Torus

In this chapter we present some ideas for extending our results from the previous chapter to the q-torus T^q. The dual group is $\mathbb{Z}^q$; this is a special case of the general fact that the dual group of a product is the product of the dual groups.

We denote a point on T^q by $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_q)$ where $-\pi \leq \phi_i < \pi$ for $i = 1, \ldots, q$; similarly, a point in $\mathbb{Z}^q$ will be represented by $\mathbf{k} = (k_1, \ldots, k_q)$. The notation $\mathbf{k} \cdot \boldsymbol{\phi}$ will represent $k_1\phi_1 + \cdots + k_q\phi_q$. We will use dm_q to represent normalized Lebesgue measure on T^q.

First, let us see what the Fourier transform on T^q does. Let $f : T^q \to \mathbb{C}$ with $f \in L^1(T^q)$. Then the Fourier coefficients of f are

$$c_{\mathbf{k}} = \hat{f}(\mathbf{k}) = \int_{T^q} f(\boldsymbol{\phi}) e^{-i\mathbf{k}\cdot\boldsymbol{\phi}} dm_q$$
$$= \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(\phi_1, \ldots, \phi_q) e^{-ik_1\phi_1} \ldots e^{-ik_2\phi_2} d\phi_1 \ldots d\phi_m.$$

Similarly, if μ is a finite regular measure on T^q, its Fourier-Stieltjes transform is

$$c_{\mathbf{k}} = \hat{\mu}(\mathbf{k}) = \int_{T^q} e^{-i\mathbf{k}\cdot\boldsymbol{\phi}}\, d\mu(\boldsymbol{\phi})$$

The infinite Toeplitz operator generated by μ is the linear operator T_μ on $L^2(\mathbb{Z}^q)$ given by

$$(T_\mu g)(\mathbf{k}) = \sum_{\mathbf{j}} \hat{\mu}(\mathbf{k} - \mathbf{j}) g(\mathbf{j}).$$

In keeping with our general approach, we plan to approximate the infinite Toeplitz operator with a sequence of finite Toeplitz operators, formed by integrating over compact subsets of the dual group. In this case, that means summing over a finite number of points in $\mathbb{Z}^q$. We will then try to show that these finite Toeplitz operators have a limiting eigenvalue distribution analogous to the one-variable case. We will explore two ideas for how to carry this out.

3.1 First Idea: Bijection from $\mathbb{Z}^q$ to $\mathbb{Z}$

In this section we consider a sequence of finite Toeplitz operators generated by a bijection between $\mathbb{Z}^q$ and $\mathbb{Z}$; that is, we use any sequence of sets $E_1 \subset E_2 \subset \dots$ such that E_n contains n points. The motivation behind this approach is to mimic our proof on the circle: if we can say something about the limiting ratio of determinants, this will allow us to find the average logarithm of the eigenvalues, and we can proceed as before.

Since Szegö's theorem was crucial to our success on the unit circle, we are very interested in extending it to the q-torus if possible. We will attempt to follow the same procedure as before.

Definition 4. *Let A denote the set of continuous $f : T^q \to \mathbb{C}$ such that $\hat{f}(\mathbf{k}) = 0$ for all $\mathbf{k} \notin \mathbb{Z}^q_+$, and $A_0 = \{f \in A \mid \int_{T^q} f dm_q = 0\}$.*

As before, the Poisson integrals allows us to identify each $f \in A$ with an analytic function on D^q:

$$\tilde{f}(\mathbf{z}) = \int_{T^q} P(\mathbf{z}, \mathbf{x}) f(\mathbf{x}) \, dm_q(\mathbf{w}).$$

Here $P(\mathbf{z}, \mathbf{w})$ is the Poisson kernel $P_{r_1}(\theta_1 - \phi_1) \dots P_{r_q}(\theta_q - \phi_q)$ where $z_j = r_j e^{i\theta_j}$ and $w_j = e^{i\phi_j}$ for $j = 1, \dots, q$, and $P_r(\theta) = \frac{1-r^2}{1-2r\cos(\theta)+r^2}$ is the one-dimensional Poisson kernel. See Appendix C for more..

Note that $\tilde{f}(0) = \int_{T^q} f \, dm_q$ so that A_0 corresponds to functions which vanish at the origin.

Unfortunately it is no longer true that the real parts of functions in A_0 are uniformly dense in $C(T^q)$. To see this, consider the 2-dimensional case. Nothing in A_0 can have a real part uniformly close to $\cos(\theta - \phi)$ because that would make its $(1, -1)$ Fourier coefficient close to $\frac{1}{2}$ and therefore nonzero. The problem here is that the trigonometric polynomials in A_0 correspond only to the first quadrant of the dual group $\mathbb{Z}^2$, that is, pairs of integers which are both positive; the real parts of trigonometric polynomials in A_0 will then include terms from the first and third quadrants. A trigonometric polynomial generated from the second and fourth quadrants, such as $\cos(\theta - \phi)$, will not be approximable.

To fix this problem, we could try to expand our definition of A_0. We will lose the connection with analytic functions which was so useful in the one-variable case, since 41 in Appendix C says that boundary values of analytic functions are precisely those whose Fourier coefficients vanish outside the first hyperquadrant. However, we may be able to salvage this approach. Recall that the relevant point at which the analyticity of the Poisson integral was invoked was in proving that $e^g - 1 \in A_0$ if $g \in A_0$. If we could find another way to prove this, we could still get Szegö's theorem on the torus.

The problem seemed to be that A_0 and its conjugates only covered some of $\mathbb{Z}^q$. Suppose we define A to include all functions f such that $\hat{f}(n_1, \ldots, n_q) = 0$ whenever $n_1 < 0$. That is, we only restrict the Fourier transform with respect to the first component in $\mathbb{Z}^q$, so that we are now sweeping out half of $\mathbb{Z}^q$ (and the conjugates of functions in A_0 will sweep out the other half).

Unfortunately we have not been able to determine whether this leads anywhere. We investigated some special cases, but Maple did not cooperate with the integrals involved. As an example of the type of question being asked, suppose $f(\theta, \phi) = e^{i(\theta-\phi)} + e^{i(5\theta+3\phi)} - 5e^{i(2\theta-7\phi)}$. We want to know whether

$$\hat{e^f}(-2,2) = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{e^{i(\theta-\phi)}+e^{i(5\theta+3\phi)}-5e^{(2\theta-2\phi)}} e^{i(2\theta-2\phi)}d\theta d\phi$$

is equal to zero.

Still more unfortunately, the bijection approach is doomed even if Szegö's theorem still holds. Recall that in order to connect Szegö's theorem to Toeplitz matrices, we had to show that the minimum of the quadratic form Q_n with the restriction $u_0 = 1$ was $\frac{D_n}{D_{n-1}}$. We did this by showing that the minimum with the restriction $u_n = 1$ was $\frac{D_n}{D_{n-1}}$, and then proving that these two constraints gave the same minimum. In higher dimensions these constraints are no longer equivalent; in fact, the minimum subject to $u_n = 1$ is still $\frac{D_n}{D_{n-1}}$, but the minimum subject to $u_0 = 1$ generally is not. Recall that we got from the constraint $u_n = 1$ to $u_0 = 1$ by conjugating an expression (which does not change its absolute value) and then multiplying it by a complex number of unit length. For that to work in higher dimensions, we would need the shape of the support of our operator (i.e. the set of points in $\mathbb{Z}^q$ used to generate it) to be such that a reflection through the origin and a translation would bring it back to its original position. To achieve this while only adding one point at a time, we must step by the same amount in the same direction each time we add a point, so that we will only cover a line instead of all $\mathbb{Z}^q$. To sum up, our problem is that we had a bijection from $\mathbb{Z}^q$ to $\mathbb{Z}$, but we actually needed an isomorphism (which of course does not exist).

Illustrative Example

Here's a numerical example to illustrate this:

Consider the measure μ on T^2 with $d\mu = \theta^2 cosh(\phi) dm(\theta) dm(\phi)$ where we have used θ and ϕ instead of ϕ_1 and ϕ_2 as our two angle coordinates. We begin our enumeration of $\mathbb{Z}^2$ with the points $(0,0), (0,1), (1,0),$ and $(0,2)$.

First let's compute some Fourier coefficients:

$$\begin{aligned}
c_{0,0} &= & \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \theta^2 cosh(\phi)\, d\theta d\phi &= & 12.09381 & \\
c_{0,1} &= & \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-i\phi}\theta^2 cosh(\phi)\, d\theta d\phi &= & -6.04691 &= c_{0,-1} \\
c_{1,0} &= & \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-i\theta}\theta^2 cosh(\phi)\, d\theta d\phi &= & -7.35216 &= c_{-1,0} \\
c_{0,2} &= & \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-2i\phi}\theta^2 cosh(\phi)\, d\theta d\phi &= & 2.41876 &= c_{0,-2} \\
c_{1,-1} &= & \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-i\theta}e^{-i\phi}\theta^2 cosh(\phi)\, d\theta d\phi &= & 3.67608 &= c_{-1,1} \\
c_{-1,2} &= & \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{i\theta}e^{-2i\phi}\theta^2 cosh(\phi)\, d\theta d\phi &= & -1.47043 &= c_{1,-2}
\end{aligned}$$

The first four Toeplitz matrices are

$$\begin{aligned}
T_0 = \begin{bmatrix} c_{0,0} \end{bmatrix} &= \begin{bmatrix} 12.09381 \end{bmatrix} \\
T_1 = \begin{bmatrix} c_{0,0} & c_{0,1} \\ c_{0,-1} & c_{0,0} \end{bmatrix} &= \begin{bmatrix} 12.09381 & -6.04691 \\ -6.04691 & 12.09381 \end{bmatrix} \\
T_2 = \begin{bmatrix} c_{0,0} & c_{0,1} & c_{1,0} \\ c_{-,-1} & c_{0,0} & c_{1,-1} \\ c_{-1,0} & c_{-1,1} & c_{0,0} \end{bmatrix} &= \begin{bmatrix} 12.09381 & -6.04691 & -7.35216 \\ -6.04691 & 12.09381 & 3.67608 \\ -7.35216 & 3.67608 & 12.09381 \end{bmatrix} \\
T_3 = \begin{bmatrix} c_{0,0} & c_{0,1} & c_{1,0} & c_{0,2} \\ c_{0,-1} & c_{0,0} & c_{1,-1} & c_{0,1} \\ c_{-1,0} & c_{-1,1} & c_{0,0} & c_{-1,2} \\ c_{0,-2} & c_{0,-1} & c_{1,-2} & c_{0,0} \end{bmatrix} &= \begin{bmatrix} 12.09381 & -6.04691 & -7.35216 & 2.41876 \\ -6.04691 & 12.09381 & 3.67608 & -6.04691 \\ -7.35216 & 3.67608 & 12.09381 & -1.47043 \\ 2.41876 & -6.04691 & -1.47043 & 12.09381 \end{bmatrix}
\end{aligned}$$

with determinants $D_0 = 12.09381,\ D_1 = 109.6951,\ D_2 = 836.3408,\ D_3 = 7552.19$.

For a polynomial $g(y,z) = u_0 + u_1 z + u_2 y + u_3 z^2$,

$$\langle g, g\rangle = \int_{T^2} g(e^{i\theta}, e^{i\phi})\overline{g(e^{i\theta}, e^{i\phi})}\, d\mu = \begin{pmatrix} \bar{u_0} & \bar{u_1} & \bar{u_2} & \bar{u_3}\end{pmatrix} T_3 \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

Now let's find some orthogonal polynomials. Using our formula for the Gram-Schmidt process,

$$\begin{aligned}
\psi_0(y,z) &= \frac{1}{\sqrt{D_0}}\left|1\right| = 0.28755 \\
\psi_1(y,z) &= \frac{1}{\sqrt{D_1 D_0}}\begin{vmatrix} c_{0,0} & c_{0,1} \\ 1 & z \end{vmatrix} = 0.33214z + 0.16602 \\
\psi_2(y,z) &= \frac{1}{\sqrt{D_2 D_1}}\begin{vmatrix} c_{0,0} & c_{0,1} & c_{1,0} \\ c_{0,-1} & c_{0,0} & c_{1,-1} \\ 1 & z & y \end{vmatrix} = 0.36216z + 0.22017 \\
\psi_3(y,z) &= \frac{1}{\sqrt{D_3 D_2}}\begin{vmatrix} c_{0,0} & c_{0,1} & c_{1,0} & c_{0,2} \\ c_{0,-1} & c_{0,0} & c_{1,-1} & c_{0,1} \\ c_{-1,0} & c_{-1,1} & c_{0,0} & c_{-1,2} \\ 1 & z & y & z^2 \end{vmatrix} = 0.33278y^2 - 0.022186
\end{aligned}$$

Now, the polynomial

$$g(y,z) = \frac{0.28755}{A}\psi_0(y,z) + \frac{0.16602}{A}\psi_1(y,z) + \frac{0.22017}{A}\psi_2(y,z) + \frac{-0.022186}{A}\psi_3(y,z),$$

where $A = 0.28755^2 + 0.16602^2 + 0.22017^2 + 0.022186^2$, has a constant term $u_0 = 1$; we also have by orthonormality

$$\langle g, g\rangle = \left(\frac{0.28755}{A}\right)^2 + \left(\frac{0.16602}{A}\right)^2 + \left(\frac{0.22017}{A}\right)^2 + \left(\frac{-0.022186}{A}\right)^2 = \frac{1}{A} = 6.28083.$$

However, note that $\frac{D_3}{D_2} = 9.03004$ which is clearly not equal to the minimum length with $u_0 = 1$.

3.2 Second Approach: Product Measures

Another idea is to consider the special case of product measures, i.e. measures of the form $\mu = \mu_1 \times \cdots \times \mu_q$ where the μ_k are measures on T. Then

$$\begin{aligned}\hat{\mu}(\mathbf{k}) &= \int_{T^q} e^{-i(k_1\phi_1+\cdots+k_q\phi_q)}\, d(\mu_1 \times \cdots \times \mu_q) \\ &= \left(\int_T e^{-ik_1\phi_1}\, d\mu_1(\phi_1)\right) \cdots \left(\int_T e^{-ik_2\phi_2}\, d\mu_2(\phi_2)\right) = \hat{\mu_1}(k_1)\cdots\hat{\mu_q}(k_q).\end{aligned}$$

Recall that the Toeplitz operator generated by a measure on T^q is

$$(T_\mu g)(\mathbf{k}) = \sum_{\mathbf{j}} \hat{\mu}(\mathbf{k}-\mathbf{j})g(\mathbf{j}).$$

The range of summation for $\mathbf{j}$ depends on how we choose to construct our finite Toeplitz operators. If we form them by summing over rectangular regions in $\mathbb{Z}^q$, then we will be able to factor the expression for $T_\mu g$. Suppose then that $\mathbf{j}$ is summed over the rectangle $[0,n]^q$, and consider first the case where the vector $g(\mathbf{j})$ is a tensor product $g = g_1 \otimes \cdots \otimes g_q$ so that $g(\mathbf{j}) = g_1(j_1)\cdots g_q(j_q)$. Then

$$\begin{aligned}(T_\mu g)(\mathbf{k}) &= \sum_{j_1,\ldots,j_q} \hat{\mu_1}(k_1-j_1)\cdots\hat{\mu_q}(k_q-j_q)g_1(j_1)\cdots g_q(j_q) \\ &= \left(\sum_{j_1}\hat{\mu_1}(k_1-j_1)g_1(j_1)\right)\cdots\left(\sum_{j_q}\hat{\mu_q}(k_q-j_q)g_q(j_q)\right) \\ &= \Big((T_{\mu_1}(g_1))(k_1)\Big)\cdots\Big((T_{\mu_q}(j_q))(k_q)\Big).\end{aligned}$$

Thus, $T_\mu = T_{\mu_1} \otimes \cdots \otimes T_{\mu_q}$. In particular, if $\mathbf{x_1}, \cdots, \mathbf{x_q}$ are eigenvectors of $T_{\mu_1}, \cdots, T_{\mu_q}$ respectively (recall that Toeplitz operators are Hermitian so that we are guaranteed a basis of eigenvectors), with corresponding eigenvalues $\lambda_1, \cdots, \lambda_q$, then

$$\begin{aligned}T_\mu(\mathbf{x_1} \otimes \cdots \otimes \mathbf{x_q}) &= (T_{\mu_1}(\mathbf{x_1})) \otimes \cdots \otimes (T_{\mu_q}(\mathbf{x_q})) \\ &= (\lambda_1\mathbf{x_1}) \otimes \cdots \otimes (\lambda_q\mathbf{x_q}) = (\lambda_1\cdots\lambda_q)\mathbf{x_1} \otimes \cdots \otimes \mathbf{x_q}\end{aligned}$$

so that $\mathbf{x_1} \otimes \cdots \otimes \mathbf{x_q}$ is an eigenvector of T_μ with eigenvalues $\lambda_1\cdots\lambda_q$. This gives us $(n+1)^q$ independent eigenvectors for T_μ; since T_μ acts on an $(n+1)^q$-dimensional space, this implies that the eigenvalues of T_μ are precisely the products of the form $(\lambda_1\cdots\lambda_q)$ where λ_k is an eigenvalue of T_{μ_k}.

Intuitively, if the distributions of the eigenvalues for each one-dimensional operator are converging, in some sense, to the derivative of the corresponding measure, then the products of those eigenvalues should be converging to the product of the derivatives, which is just the derivative of the product. (An example to clarify this last statement: If $\mu = \mu_1 \times \mu_2 = (\mu_{1a} + \mu_{1s}) \times (\mu_{2a} + \mu_{2s}) = \mu_{1a}\mu_{2a} + \mu_{1a}\mu_{2s} + \mu_{2a}\mu_{2s} + \mu_{1s}\mu_{2s}$ where μ_{1a} and μ_{2a} are absolutely continuous and μ_{1s} and μ_{2s} are singular, then $\mu_{1a}\mu_{2s}, \mu_{1s}\mu_{2a}$, and $\mu_{1s}\mu_{2s}$ are all singular, and the absolutely continuous part of $\mu_1 \times \mu_2$ is $\mu_{1a}\mu_{2a}$.)

We will now prove a more precise statement of this, using the two-dimensional case to illustrate a general concept.

Lemma 1. *Let $\epsilon > 0$, let $U \subset \mathbb{R}$ be a finite union of open intervals, and let μ be a positive measure on $\mathbb{R}$ such that $\mu(\mathbb{R}) = 1$ and any point $x \in \mathbb{R} \setminus U$ has $\mu(\{x\}) < \epsilon$. Then $\mathbb{R} \setminus U$ may be partitioned into finitely many closed intervals $[x_j, y_j]$ (possibly overlapping at the endpoints) with $\mu([x_j, y_j]) < 2\epsilon$. Additionally, suppose $\alpha \in \mathbb{R}$ and σ is another positive measure with $\sigma(\mathbb{R}) = 1$ and $\sigma(\{x\}) = 0$ for x in the boundary of U; if $\alpha = 0$ we also require $\sigma(\{0\}) = 0$. Then x_j and y_j may be chosen such that $x_j, y_j \neq 0$ and $\mu(\{x_j\}) = \sigma\left(\left\{\frac{\alpha}{x_j}\right\}\right) = \mu(\{y_j\}) = \sigma\left(\left\{\frac{\alpha}{y_j}\right\}\right) = 0$.*

Proof. $\mathbb{R} \setminus U$ is a union of finitely many closed intervals, so it is sufficient to give the desired construction on one of them. Suppose $I = (-\infty, a]$ is one such interval; the construction will be quite analogous for $[b, \infty)$ or for a finite interval. Let $x_1 = \sup\{x \in I \mid \mu((-\infty, x]) < \epsilon\}$. (To see that this set is nonempty, consider that

$$\bigcap_{\substack{n \in \mathbb{Z} \\ -n \leq a}} (-\infty, -n] = \emptyset$$

so that $\lim_{n\to\infty} \mu((-\infty, -n]) = 0$.) Then if $x_1 < a$ we have $\epsilon \leq \mu((-\infty, x_1]) < 2\epsilon$, because

$$\mu((-\infty, x_1)) = \mu\left(\bigcup_{n=1}^{\infty}\left(-\infty, x_1 - \frac{1}{n}\right]\right) = \lim_{n\to\infty} \mu\left(\left(-\infty, x_1 - \frac{1}{n}\right]\right) \leq \epsilon$$

so that $\mu((-\infty, x_1]) = \mu((-\infty, x_1)) + \mu(\{x_1\}) < 2\epsilon$, and similarly,

$$\mu((-\infty, x_1]) = \mu\left(\bigcap_{n=1}^{\infty}\left(-\infty, x_1 + \frac{1}{n}\right]\right) = \lim_{n\to\infty} \mu\left(\left(-\infty, x_1 + \frac{1}{n}\right]\right) \geq \epsilon.$$

(If $x_1 = a$ we will still have $\mu((-\infty, x_1]) < 2\epsilon$ but may have $\mu((-\infty, x_1]) < \epsilon$.) Using continuity of measures again (and the finiteness of μ), $\epsilon \leq \mu((-\infty, x]) < 2\epsilon$ for x in some interval to the right of x_1 (again, assuming $x_1 < a$); we may move x_1 to some point in this interval so as to guarantee $\mu(\{x_1\}) = \sigma\left(\left\{\frac{\alpha}{x_1}\right\}\right) = 0$. (In so doing, we only need to avoid the countably many points z for which $\mu(\{z\}) \neq 0$ or $\sigma\left(\left\{\frac{\alpha}{z}\right\}\right) \neq 0$, so this is clearly possible.)

We have now constructed the first of our subintervals. If $x_1 < a$ we let $x_2 = \sup\{x \in I \mid \mu([x_1, x]) < \epsilon\}$; then $\epsilon \leq \mu([x_1, x_2] < 2\epsilon$ by similar reasoning, and once again we can adjust x_1 slightly to the right if needed. We continue this process until some x_k is equal to a; this happens in finitely many steps because μ is a finite positive measure and each interval has measure at least ϵ. We apply the same construction to each of the finitely many closed intervals that make up $\mathbb{R} \setminus U$. $\square$

Theorem 10. *Let μ_1 and μ_2 be finite signed measures on T, with derivatives f and g wrt m, and $\lambda_k^{(n)}$ and $\nu_k^{(n)}$ the eigenvalues of their respective Toeplitz operators. Define m_f and m_g on Borel subsets $E \subset \mathbb{R}$ by $m_f(E) = m(f^{-1}(E))$ and $m_g(E) = m(g^{-1}(E))$. Let $\alpha \in \mathbb{R}$ such that $m_2(\{(x,y) \mid f(x)g(y) = \alpha\}) = 0$, where m_2 denotes normalized Lebesgue measure on T^2. Then*

$$\lim_{n\to\infty} \frac{\#\text{ of products } \lambda_k^{(n)}\nu_j^{(n)} \text{ in } [\alpha, \infty)}{(n+1)^2} = m_2(\{(x,y) \mid f(x)g(y) \geq \alpha\}).$$

Proof. We give the proof first for $\alpha \neq 0$. The case $\alpha = 0$ is more complicated and will be discussed at the end of the section.

Our strategy will be as follows: We want to take two copies of $\mathbb{R}$ and divide both into "small" intervals so as to have control over the products taken from pairs of these intervals; our notion of "small" will depend on m_f and m_g. Our partition of the first line will determine that of the second. On the first one, we will construct small intervals around troublesome points (these being zero and those points x for which $m_f(\{x\})$ is greater than some fixed amount), and apply the construction from the lemma to the remainder of the line.

Let $\epsilon > 0$. Let $c_1, \ldots, c_N$ be the nonzero points for which $m_f(\{c_k\}) > \frac{\epsilon}{12}$, and $c_0 = 0$ (whether or not $m_f(\{0\}) > \frac{\epsilon}{12}$). By the continuity of m_g, we can construct open intervals $A_k = (a_k, b_k)$ for $k = 0, \ldots, n$ such that

- $c_k \in A_k$
- $A_k \cup A_j = \emptyset$ for $j \neq k$

- $m_g((\frac{\alpha}{b_k}, \frac{\alpha}{a_k})) < \frac{\epsilon}{6}$ or $m_g((\frac{\alpha}{a_k}, \frac{\alpha}{b_k})) < \frac{\epsilon}{6}$ (whichever is the correct ordering of the endpoints, depending on the signs of $\alpha, a_k,$ and b_k)
- $m_f(\{a_k\}) = m_f(\{b_k\}) = 0$.

For the last condition, note that there are only countably many points we need to avoid, so we can easily construct our intervals with none of them as endpoints. For convenience, we will also stipulate that $|a_0| = |b_0|$, so we may write $(a_0, b_0) = (-d, d)$. By Lemma (1), we can divide $\mathbb{R} \setminus \bigcup_k A_k$ into finitely many intervals $[x_j, y_j]$ for $j = 1, \ldots, M$ (sometimes overlapping at the endpoints) with $m_f([x_j, y_j]) < \frac{\epsilon}{6}$ and $m_f(\{x_j\}) = m_f(\{y_j\}) = m_g\left(\left\{\frac{\alpha}{x_j}\right\}\right) = m_g\left(\left\{\frac{\alpha}{y_j}\right\}\right) = 0$.

We now define intervals $I_k, J_k, \tilde{I}_j, \tilde{J}_j$ as follows:

$$
\begin{aligned}
I_0 &= \mathbb{R} \\
J_0 &= \emptyset \\
I_k &= \begin{cases} \left(\frac{\alpha}{b_k}, \infty\right) & \text{if } a_k, b_k > 0 \\ \left(-\infty, \frac{\alpha}{a_k}\right) & \text{if } a_k, b_k < 0 \end{cases} \\
J_k &= \begin{cases} \left(\frac{\alpha}{a_k}, \infty\right) & \text{if } a_k, b_k > 0 \\ \left(-\infty, \frac{\alpha}{b_k}\right) & \text{if } a_k, b_k < 0 \end{cases} \\
\tilde{I}_j &= \begin{cases} \left[\frac{\alpha}{y_j}, \infty\right) & \text{if } x_j, y_j > 0 \\ \left(-\infty, \frac{\alpha}{x_j}\right] & \text{if } x_j, y_j < 0 \end{cases} \\
\tilde{J}_j &= \begin{cases} \left[\frac{\alpha}{x_j}, \infty\right) & \text{if } x_j, y_j > 0 \\ \left(-\infty, \frac{\alpha}{y_j}\right] & \text{if } x_j, y_j < 0 \end{cases}
\end{aligned}
$$

where k is understood to refer to the case $k > 0$. The point of these constructions is that $x \in (a_k, b_k)$ and $y \in J_k \implies xy \geq \alpha$ whereas $x \in (a_k, b_k)$ and $xy \geq \alpha \implies y \in I_k$, so that

$$(a_k, b_k) \times J_k \subset \{(x, y) \mid x \in (a_k, b_k) \text{ and } xy \geq \alpha\} \subset (a_k, b_k) \times I_k.$$

The analogous statement holds for $[x_j, y_j]$, $\tilde{I}_j$, and $\tilde{J}_j$. Also note that the sets $I_k \setminus J_k$ and $\tilde{I}_j \setminus \tilde{J}_j$ are all disjoint (excluding $k = 0$) and that $m_g(\{y\}) = 0$ for y in the boundary of $I_k, J_k, \tilde{I}_j$, or $\tilde{J}_j$.

We introduce the notation $\Phi_n(f,E) = \frac{\#\text{ of }\lambda_i^{(n)}\text{ in }E}{n+1}$ for any $E \subset \mathbb{R}$, and similarly for $\Phi_n(g,E)$. Then we know from the one-variable case that

$$\lim_{n\to\infty} \Phi_n(f,A) = m_f(A)$$

and

$$\lim_{n\to\infty} \Phi_n(g,B) = m_g(B)$$

where A is any one of the intervals (a_k,b_k) or $[x_j,y_j]$ and B is any one of the I_k, J_k, $\tilde{I}_j$, or $\tilde{J}_j$. There are finitely many of the A's and B's, so for any $\delta > 0$, we can choose n sufficiently large that $|\Phi_n(f,A) - m_f(A)| < \delta$ for all of the A's and $|\Phi_n(g,B) - m_g(B)| < \delta$ for all the B's. Then

$$\begin{aligned}
&\frac{\#\text{ of products}\lambda_i^{(n)}\nu_l^{(n)}\text{ in }[\alpha,\infty)}{(n+1)^2}\\
&\leq \sum_k \left(\frac{\#\text{ of }\lambda_i^{(n)}\text{ in }(a_k,b_k)}{n+1}\right)\left(\frac{\#\text{ of }\nu_l^{(n)}\text{ in }I_k}{n+1}\right)\\
&+\sum_j \left(\frac{\#\text{ of }\lambda_i^{(n)}\text{ in }[x_j,y_j]}{n+1}\right)\left(\frac{\#\text{ of }\nu_l^{(n)}\text{ in }\tilde{I}_j}{n+1}\right)\\
&\leq \sum_k \left(m_f((a_k,b_k))+\delta\right)\left(m_g(I_k)+\delta\right)+\sum_j\left(m_f([x_j,y_j])+\delta\right)\left(m_g(\tilde{I}_j)+\delta\right)\\
&\qquad \leq \sum_k m_f((a_k,b_k))m_g(I_k)+\sum_j m_f([x_j,y_j])m_g(\tilde{I}_j)+2\delta+\delta^2(M+N).
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\frac{\#\text{ of products }\lambda_i^{(n)}\nu_l^{(n)}\text{ in }[\alpha,\infty)}{(n+1)^2}\\
&\geq \sum_k \left(\frac{\#\text{ of }\lambda_i^{(n)}\text{ in }(a_k,b_k)}{n+1}\right)\left(\frac{\#\text{ of }\nu_l^{(n)}\text{ in }J_k}{n+1}\right)\\
&+\sum_j \left(\frac{\#\text{ of }\lambda_i^{(n)}\text{ in }[x_j,y_j]}{n+1}\right)\left(\#\text{ of }\nu_l^{(n)}\text{ in }\tilde{J}_j\right)\\
&\geq \sum_k \left(\left(m_f((a_k,b_k))-\delta\right)\left(m_g(J_k)-\delta\right)-\delta^2\right)+\sum_j\left(\left(m_f([x_j,y_j])-\delta\right)\left(m_g\left(\tilde{J}_j\right)-\delta\right)-\delta^2\right)\\
&\qquad \geq \sum_k m_f((a_k,b_k))m_g\left(J_k\right)+\sum_j m_f([x_j,y_j])m_g\left(\tilde{J}_j\right)-2\delta.
\end{aligned}$$

(Here the subtraction of δ^2 in the fourth line is necessary because of the possibility that, for example, $m_f((a_k, b_k)$ and $m_g(J_k)$ are both less than δ.) Now we take δ small enough that $2\delta + \delta^2(M+N) < \frac{\epsilon}{4}$; then we have, for sufficiently large n,

$$\begin{aligned}
&\sum_k m_f((a_k, b_k)) m_g\left(J_k\right) + \sum_j m_f([x_j, y_j]) m_g\left(\tilde{J}_j\right) - \frac{\epsilon}{4} \\
&\leq \frac{\#\text{ of products } \lambda_i^{(n)} \nu_l^{(n)} \text{ in } [\alpha, \infty)}{(n+1)^2} \\
&\qquad\qquad \leq \sum_k m_f((a_k, b_k)) m_g\left(I_k\right) + \sum_j m_f([x_j, y_j]) m_g\left(\tilde{J}_j\right) + \frac{\epsilon}{4}.
\end{aligned}$$

But it is also clear that

$$\begin{aligned}
&\sum_k m_f((a_k, b_k)) m_g\left(J_k\right) + \sum_j m_f([x_j, y_j]) m_g\left(\tilde{J}_j\right) \\
&\leq m_2\left(\{(x, y) \mid f(x) g(y) \geq \alpha\}\right) \\
&\qquad\qquad \leq \sum_k m_f((a_k, b_k)) m_g\left(I_k\right) + \sum_j m_f([x_j, y_j]) m_g\left(\tilde{I}_j\right)
\end{aligned}$$

so certainly

$$\begin{aligned}
&\sum_k m_f((a_k, b_k)) m_g\left(J_k\right) + \sum_j m_f([x_j, y_j]) m_g\left(\tilde{J}_j\right) - \frac{\epsilon}{4} \\
&\leq m_2\left(\{(x, y) \mid f(x) g(y) \geq \alpha\}\right) \\
&\qquad\qquad \leq \sum_k m_f((a_k, b_k)) m_g\left(I_k\right) + \sum_j m_f([x_j, y_j]) m_g\left(\tilde{I}_j\right) + \frac{\epsilon}{4}.
\end{aligned}$$

The upper and lower bounds differ by

$$\begin{aligned}
&\frac{\epsilon}{2} + \sum_k m_f((a_k, b_k)) m_g\left(I_k \setminus J_k\right) + \sum_j m_f([x_j, y_j]) m_g\left(\tilde{I}_j \setminus \tilde{J}_j\right) \\
&\qquad \leq \frac{\epsilon}{2} + \sum_k \frac{\epsilon}{6} m_g\left(I_k \setminus J_k\right) + \sum_j \frac{\epsilon}{6} m_g\left(\tilde{I}_j \setminus \tilde{J}_j\right) \\
&\quad = \frac{\epsilon}{2} + \frac{\epsilon}{6} + \sum_{k \neq 0} \frac{\epsilon}{6} m_g\left(I_k \setminus J_k\right) + \sum_j \frac{\epsilon}{6} m_g\left(\tilde{I}_j \setminus \tilde{J}_j\right) \leq \frac{\epsilon}{2} + \frac{\epsilon}{6} + \frac{\epsilon}{6} + \frac{\epsilon}{6} = \epsilon
\end{aligned}$$

since $m_g(\mathbb{R}) = 1$ and $I_k \setminus J_k$, $\tilde{I}_j \setminus \tilde{J}_j$ for $k \neq 0$ are all disjoint. This shows that, for sufficiently large n,

$$\frac{\#\text{ of products } \lambda_i^{(n)} \nu_l^{(n)} \text{ in } [\alpha, \infty)}{(n+1)^2}$$

and

$$m_2\left(\{(x,y)\mid f(x)g(y)\geq\alpha\}\right)$$

are both contained in the same interval of size at most ϵ, and hence are within ϵ of each other. Thus,

$$\lim_{n\to\infty}\frac{\text{\# of products }\lambda_i^{(n)}\nu_l^{(n)}\text{ in }[\alpha,\infty)}{(n+1)^2}=m_2\left(\{(x,y)\mid f(x)g(y)\geq\alpha\}\right).$$

This finishes the proof for the case $\alpha\neq 0$. Now suppose $\alpha=0$. The condition $m_2(\{(x,y)\mid f(x)g(y)=\alpha\})=0$ implies that at least one of $m_f(\{0\})$ and $m_g(\{0\})$ is zero; WLOG, we assume $m_g(\{0\})=0$. We proceed as before, except that the third condition in the construction of the A_k intervals becomes trivial. The rest of the proof follows as before. □

To extend to higher dimensions, we can use induction and a very slight adaptation of the last theorem. For example, suppose we have measures μ_1,μ_2,μ_3 with derivatives f,g,h. Although the theorem was stated for pairs of measures on T, the same construction works if one of the pair is actually a measure on T^2. Suppose $\lambda_k^{(n)}$, $\nu_k^{(n)}$, and $\eta_k^{(n)}$ are the eigenvalues of μ_1,μ_2, and μ_3, $\xi_k^{(n)}$ the eigenvalues of $\mu_2\times\mu_3$, and $\psi=gh$ is the derivative of $\mu_2\times\mu_3$. Then we have

$$\begin{aligned}\lim_{n\to\infty}\frac{\text{\# of products }\lambda_j^{(n)}\nu_k^{(n)}\eta_l^{(n)}\text{ in }[\alpha,\infty)}{(n+1)^3}&=\lim_{n\to\infty}\frac{\text{\# of products }\lambda_k^{(n)}\xi_l^{(n)}\text{ in }[\alpha,\infty)}{(n+1)^3}\\&=m_3(\{(x,y)\mid x\in T,y\in T^2,f(x)\psi(y)\geq\alpha\})\\&=m_3(\{(x,y,z)\mid x,y,z\in T,f(x)g(y)h(z)\geq\alpha\}).\end{aligned}$$

One might wonder how proving our theorem for product measures will help prove it in general. If we specialize to absolutely continuous measures, functions of the form

$$f(\phi_1,\ldots\phi_q)=\sum_k f_{1k}(\phi_1)\ldots f_{qk}(\phi_q)$$

are uniformly dense in $C(T^q)$ by the Stone-Weierstrass theorem. This in turn implies that they are dense in L^1. This provides some hope of extending to other absolutely continuous measures. For general measures, sums of product measures are no longer dense in total variation, but perhaps some kind of weaker convergence still holds. It is an open question, and easily another entire thesis in itself, what conditions on a sequence of measures will guarantee the convergence of the eigenvalue distributions of their Toeplitz operators.

Appendix A

Fourier Transforms on Locally Compact Abelian Groups

A.1 Introduction to LCA Groups

Definition 5. *A topological group G is a group equipped with a Hausdorff topology such that $(x, y) \mapsto xy^{-1}$ is a continuous function on $G \times G$. If G is an Abelian group and its topology is locally compact, G is a locally compact Abelian (LCA) group.*

An equivalent definition of topological group is that the functions $x \mapsto x^{-1}$ and $x \mapsto xy$ are continuous on G.

LCA groups provide the framework for the abstract study of Fourier analysis. With a locally compact Hausdorff topology, many nice theorems from measure theory (e.g. the Riesz representation theorem and the Radon-Nikodym theorem) can be brought to bear. With an Abelian group structure, we have a notion of translation, which allows us to talk about convolutions; as we will see, we are also guaranteed the existence of a translation invariant measure, which makes many computations run smoothly. In short, we have all the basic ingredients we need to do Fourier analysis.

We now present without proof some basic facts about LCA groups, taken from [Rudin (1990)]. We follow the convention of writing Abelian group operations as addition.

A.2 Haar Measure and Convolution

Theorem 11. *For any LCA group G, there exists a positive regular measure m on G which is not identically zero and is translation invariant, i.e.*

$$m(E + x) = m(E)$$

for all $x \in G$ and every Borel set $E \subset G$. This measure is unique up to a multiplicative constant.

A positive regular translation invariant measure on a group is called a Haar measure; this theorem allows us to speak of *the* Haar measure on a group, up to normalization. The normalization is usually taken such that $m(G) = 1$ for compact groups and m is counting measure for discrete groups; for finite groups, either one may be used.

Unless otherwise notated, all integration will be done with respect to Haar measure; dx, dy, etc. will represent Haar measure, and L^p spaces will be with respect to Haar measure. Note that on the real line, Haar measure corresponds to Lebesgue measure.

Theorem 12. *Let $f_x(y) = f(y - x)$. Then $\|f_x\|_p = \|f\|_p$ for any $f \in L^p(G)$, $1 \leq p \leq \infty$.*

Theorem 13. *The mapping $x \mapsto f_x$ is a uniformly continuous map of G into $C_0(G)$ and into $L^p(G)$ for $1 \leq p < \infty$.*

Definition 6. *For any two Borel functions f and g on G such that*

$$\int_G |f(x - y)g(y)|\, dy < \infty, \tag{A.1}$$

we define the convolution

$$(f * g)(x) = \int_G f(x - y)g(y)\, dy.$$

The condition (A.1) allows the use of Fubini's theorem, which is indispensable in proving theorems about convolutions.

Theorem 14.

- *If (A.1) holds for some $x \in G$, then $(f * g)(x) = (g * f)(x)$.*
- *If $f \in L^1(G)$ and $g \in L^\infty(G)$, then $f * g$ is bounded and uniformly continuous.*
- *If $f, g \in C_c(G)$, $f * g \in C_c(G)$.*

- *If* $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p < \infty$, $f \in L^p(G)$, *and* $g \in L^q(G)$, *then* $f * g \in C_0(G)$.
- *If* $f, g \in L^1(G)$, *then (A.1) holds for almost all* $x \in G$, $f * g \in L^1(G)$, *and*
$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$
- *If* $f, g, h \in L^1(G)$, *then* $(f * g) * h = f * (g * h)$.

The last three properties may be summarized by saying that convolution is a multiplication on $L^1(G)$ which turns it into a (commutative) Banach algebra. If G is discrete, $L^1(G)$ has a unit, namely the "delta function" defined by $e(0) = 1$ and $e(x) = 0$ for $x \neq 0$. If G is not discrete, we shall see later that $L^1(G)$ does not have a unit; however, it does have approximate units, in the sense given by the following theorem.

Theorem 15. *Given* $f \in L^1(G)$ *and* $\epsilon > 0$, *there exists a neighborhood* V *of 0 in* G *such that for any nonnegative Borel function* u *which vanishes outside* V *and for which* $\int_G u(x)\,dx = 1$,
$$\|f - f * u\|_1 < \epsilon.$$

A.3 The Dual Group and the Fourier Transform

Definition 7. *Let* G *be an LCA group. Then the dual group of* G *is the multiplicative group* Γ *of continuous characters of* G, *i.e. continuous functions* $\gamma : G \to \mathbb{C}$ *such that* $|\gamma(x)| = 1$ *for all* x *and*
$$\gamma(x + y) = \gamma(x)\gamma(y) \qquad \text{for all } x, y \in G.$$

Definition 8. *For* $f \in L^1(G)$, *the function* $\hat{f}$ *on* Γ *defined by*
$$\hat{f}(\gamma) = \int_G \overline{\gamma(x)} f(x)\,dx$$
is called the Fourier transform of f. *The set of such* $\hat{f}$ *will be denoted* $A(\Gamma)$.

Theorem 16. *Let* Γ *be given the weak topology induced by the Fourier transforms of all the functions in* $L^1(G)$. *Then* Γ *is an LCA group whose dual is homeomorphically isomorphic to* G.

The latter half of this theorem is usually thought of as saying that "the dual of the dual is the original group." To reflect this duality, the notation (x, γ) is often used instead of $\gamma(x)$; the idea is that the groups are on equal footing, with an element from either defining a function on the other.

Theorem 17.
- *$A(\Gamma)$ is a separating self-adjoint subalgebra of $C_0(\Gamma)$, and therefore is dense in $C_0(\Gamma)$.*
- $\widehat{f * g} = \hat{f}\hat{g}$.
- *$A(\Gamma)$ is invariant under translation or under multiplication by (x, γ) for a fixed $x \in G$.*
- *$\|\hat{f}\|_\infty \leq \|f\|_1$ so that the Fourier transform is a continuous map of $L^1(G)$ into $C_0(\Gamma)$.*
- *For $f \in L^1(G)$ and $\gamma \in \Gamma$, $(f * \gamma)(x) = (x, \gamma)\hat{f}(\gamma)$.*

Theorem 18. *If G is discrete, Γ is compact; if G is compact, Γ is discrete.*

A.4 Fourier-Stieltjes Transforms

Let $M(G)$ denote the set of complex regular measures on G.

Definition 9. *Let $\mu, \lambda \in M(G)$. We define a measure $\mu * \lambda$ on G by*

$$(\mu * \lambda)(E) = (\mu \times \lambda)(\{(x, y) \in G \times G \mid x + y \in E\}).$$

We call $\mu * \lambda$ the convolution of μ and λ.

Theorem 19.
1. *If $\mu, \lambda \in M(G)$, then $\mu * \lambda \in M(G)$.*
2. *For $\mu, \lambda, \nu \in M(G)$, $\mu * \lambda = \lambda * \mu$ and $(\mu * \lambda) * \nu = \mu * (\lambda * \nu)$.*
3. $\|\mu * \lambda\| \leq \|\mu\| \cdot \|\lambda\|$.

This theorem says that $M(G)$ is a commutative Banach algebra with convolution as the multiplication.

Definition 10. *For $\mu \in M(G)$, the function $\hat{\mu}$ on Γ defined by*

$$\hat{\mu}(\gamma) = \int_G (-x, \gamma)\, d\mu(x)$$

is called the Fourier-Stieltjes transform of μ (sometimes called by either name alone). The set of all $\hat{\mu}$ will be denoted $B(\Gamma)$.

Theorem 20.
- *Each $\hat{\mu} \in B(\Gamma)$ is bounded and uniformly continuous.*
- $\widehat{\mu * \lambda} = \hat{\mu}\hat{\lambda}$.

- *$B(\Gamma)$ is invariant under translation, complex conjugation, or multiplication by (x, γ) for fixed $x \in G$.*

We may identify $L^1(G)$ with the subset of $M(G)$ consisting of absolutely continuous measures. We also introduce the sets $M_c(G) \subset M(G)$ of continuous measures (ones whose values are zero on every countable set) and $M_d(G) \subset M(G)$ of discrete measures (ones that are concentrated on countable sets). The following theorem describes the place of these sets in the algebraic structure of $M(G)$.

Theorem 21. *$M_d(G)$ is a closed subalgebra of $M(G)$; $M_c(G)$ and $L^1(G)$ are closed ideals in $M(G)$.*

A.5 The Inversion Theorem

Theorem 22. *Let $B(G)$ denote the set of all functions f on G which can be represented in the form*

$$f(x) = \int_\Gamma (x, \gamma)\, d\mu(\gamma)$$

for some measure $\mu \in M(\Gamma)$.

1. *If $f \in L^1(G) \cap B(G)$, then $\hat{f} \in L^1(\Gamma)$.*
2. *For a fixed Haar measure on G, the Haar measure on Γ can be normalized such that*
   $$f(x) = \int_\Gamma \hat{f}(\gamma)(x, \gamma)\, d\gamma$$
   for every $f \in L^1(G) \cap B(G)$.

Note that we previously introduced conventions for the normalization of compact and discrete groups; it turns out that these normalizations satisfy the above condition.

A function of the form

$$f(x) = \sum_{j=1}^{n} a_j (x, \gamma_j)$$

is called a trigonometric polynomial on G. This terminology is borrowed from the cases where G is the unit circle or the real line.

Theorem 23. *If G is compact, the trigonometric polynomials on G are a dense subalgebra of $C(G)$.*

It follows from this that the trigonometric polynomials are dense in $L^p(G)$ for $1 \le p < \infty$ if G is compact.

A.6 The Plancherel Transform

Theorem 24. *The Fourier transform, restricted to $(L^1 \cap L^2)(G)$, is an isometry (with respect to L^2 norms) onto a dense linear subspace of $L^2(\Gamma)$.*

This theorem implies that the Fourier transform has a unique extension to an isometry from $L^2(G)$ to $L^2(\Gamma)$; this isometry is called the Plancherel transform and is also denoted by $f \mapsto \hat{f}$.

Theorem 25. *$A(\Gamma)$ consists precisely of convolutions $F_1 * F_2$ where $F_1, F_2 \in L^2(\Gamma)$.*

A.7 Consequences of the Duality Theorem

We have seen that Γ is an LCA group whose dual is homemorphically isomorphic to G (i.e., its dual *is* G). This has several consequences for the theory of Fourier transforms.

Theorem 26.

- *If $\mu \in M(G)$ and $\hat{\mu} = 0$ for all $\gamma \in \Gamma$, then $\mu = 0$.*
- *If G is not discrete, $L^1(G)$ has no unit. Hence $L^1(G) = M(G)$ if and only if G is discrete.*
- *If $\mu \in M(G)$ and $\hat{\mu} \in L^1(\Gamma)$, then $\exists f \in L^1(G)$ such that $d\mu(x) = f(x)\,dx$, and*
$$f(x) = \int_\Gamma \hat{\mu}(\gamma)(x, \gamma)\,d\gamma.$$

The first of these statement implies that $\hat{\mu}$ uniquely determines μ; the last says that only absolutely continuous measures have Fourier transforms in $L^1(\Gamma)$.

A.8 Characterization of $B(\Gamma)$

We introduce a norm on $B(\Gamma)$ by letting $\|\hat{\mu}\| = \|\mu\|$.

Theorem 27. *For a function $\phi : \Gamma \to \mathbb{C}$, the following are equivalent:*

- *$\phi \in B(\Gamma)$ and $\|\phi\| \leq A$.*
- *ϕ is continuous, and*
$$|\sum_{i=1}^{n} c_i \phi(\gamma_i)| \leq A\|f\|_\infty$$
for every trigonometric polynomial $f(x) = \sum_{i=1}^{n} c_i(x, \gamma_i)$.

Corollary 3. *If* $\phi_n \in B(\Gamma)$ *and* $\|\phi_n\| \leq A$ *for* $n = 1, 2, 3, \ldots$, $\phi \in C(\Gamma)$, *and* $\phi(\gamma) = \lim_{n\to\infty} \phi_n(\gamma)$, *then* $\phi \in B(\Gamma)$ *and* $\|\phi\| \leq A$.

A.9 Examples

Having defined the Fourier transform in a very general setting, we should relate it to the familiar notion of Fourier transform.

One of the familiar settings for Fourier analysis is that of periodic functions, say, functions on the reals modulo 2π. We identify $\mathbb{R}/(2\pi\mathbb{Z})$ with the unit circle T, i.e. the multiplicative group of complex numbers with unit length. It can be shown that the continuous characters on T are all of the form $\gamma(\theta) = e^{in\theta}$ for some $n \in \mathbb{Z}$. Hence, the dual group of T is $\mathbb{Z}$. The Fourier transform then becomes

$$\hat{f}(n) = \int_T \overline{e^{in\theta}} f(\theta)\, dm = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\theta} f(\theta)\, d\theta.$$

Here we have used dm for Haar measure on the circle, which is $\frac{1}{2\pi}$ times Lebesgue measure $d\theta$. We see that the familiar formula for the Fourier coefficients of a periodic function agrees with our general definition of Fourier transform. The inverse transform is given by

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{inx}.$$

Another familiar example of Fourier analysis deals with non-periodic functions on the real line. It turns out that the continuous characters of $\mathbb{R}$ are all of the form e^{ikx} for $k \in \mathbb{R}$; by the identification $k \leftrightarrow e^{ikx}$, we see that $\mathbb{R}$ is its own dual group. (The standard topology for $\mathbb{R}$ is the same as the topology it inherits as a dual group.) Then our Fourier transform becomes

$$\hat{f}(\kappa) = \int_{\mathbb{R}} \overline{e^{i\kappa x}} f(x)\, dx = \int_{-\infty}^{\infty} e^{-i\kappa x} f(x)\, dx.$$

It turns out that the normalization factor referred to in the inversion theorem is $\frac{1}{2\pi}$, so that the inverse transform is given by

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\kappa) e^{i\kappa x}\, d\kappa.$$

Of course, by choosing a different Haar measure for the "time domain" $\mathbb{R}$, the factor of $\frac{1}{2\pi}$ may appear in the transform instead of the inverse transform, or a $\frac{1}{\sqrt{2\pi}}$ may appear in both, etc.

Finally, consider the Fourier transform of a function on the integers, i.e. a sequence. It can be shown that the continuous characters on $\mathbb{Z}$ are all of the form $\gamma(n) = e^{in\alpha}$. Thus, the dual group of $\mathbb{Z}$ is the unit circle T. Of course, we already knew this because of the duality theorem. Then

$$\hat{f}(e^{i\alpha}) = \int_{\mathbb{Z}} \overline{e^{in\alpha}} f(n) dn = \sum_{-\infty}^{\infty} e^{-in\alpha} f(n)$$

since integration over the integers is just summation. The inverse transform is

$$f(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\alpha} \hat{f}(e^{in\alpha})\, d\alpha.$$

Note that the transform and inverse transform on $\mathbb{Z}$ are just the inverse transform and transform on T, respectively.

Appendix B

Formal Determinants and the Gram-Schmidt Formula

B.1 Introduction

If V is any inner product space over a field F and $\vec{v}_0, \ldots, \vec{v}_n \in V$ are independent, then it is well known that the Gram-Schmidt process guarantees the existence and uniqueness of $\vec{e}_0, \ldots, \vec{e}_n \in V$ such that

- $\langle \vec{e}_j, \vec{e}_k \rangle = \delta_{jk}$ where $\langle , \rangle$ is the inner product on V
- $\operatorname{span}(\vec{e}_0, \ldots, \vec{e}_k) = \operatorname{span}(\vec{v}_0, \ldots, \vec{v}_k)$ for $k = 0, \ldots, n$

What is less well known is an explicit formula for $\vec{e}_0, \ldots, \vec{e}_n$, namely

$$\vec{e}_k = \frac{1}{\sqrt{D_{k-1} D_k}} \begin{vmatrix} \langle \vec{v}_0, \vec{v}_0 \rangle & \langle \vec{v}_1, \vec{v}_0 \rangle & \ldots & \langle \vec{v}_k, \vec{v}_0 \rangle \\ \langle \vec{v}_0, \vec{v}_1 \rangle & \langle \vec{v}_1, \vec{v}_1 \rangle & \ldots & \langle \vec{v}_k, \vec{v}_1 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{v}_0, \vec{v}_{k-1} \rangle & \langle \vec{v}_1, \vec{v}_{k-1} \rangle & \ldots & \langle \vec{v}_k, \vec{v}_{k-1} \rangle \\ \vec{v}_0 & \vec{v}_1 & \ldots & \vec{v}_k \end{vmatrix} \tag{B.1}$$

$$\text{where } D_k = \begin{vmatrix} \langle \vec{v}_0, \vec{v}_0 \rangle & \langle \vec{v}_1, \vec{v}_0 \rangle & \ldots & \langle \vec{v}_k, \vec{v}_0 \rangle \\ \langle \vec{v}_0, \vec{v}_1 \rangle & \langle \vec{v}_1, \vec{v}_1 \rangle & \ldots & \langle \vec{v}_k, \vec{v}_1 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{v}_0, \vec{v}_{k-1} \rangle & \langle \vec{v}_1, \vec{v}_{k-1} \rangle & \ldots & \langle \vec{v}_n, \vec{v}_{k-1} \rangle \\ \langle \vec{v}_0, \vec{v}_k \rangle & \langle \vec{v}_1, \vec{v}_k \rangle & \ldots & \langle \vec{v}_k, \vec{v}_k \rangle \end{vmatrix} \neq 0$$

is the Gram determinant for $k = 0, 1, \ldots, n$, and $D_{-1} = 1$. This is what we will prove.

B.2 Formal Determinants

The term “formal determinant” refers to a “determinant” in which one of the rows consists of vectors instead of scalars. This is a notational convenience; it is defined by a cofactor expansion along the row of vectors, so

$$\begin{vmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n} \\ a_{1,0} & a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \dots & a_{n-1,n} \\ \vec{v}_0 & \vec{v}_1 & \dots & \vec{v}_n \end{vmatrix} :=$$

$$\vec{v}_0 \begin{vmatrix} a_{0,1} & a_{0,2} & \dots & a_{0,n} \\ a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,1} & \dots & a_{n-1,n} \end{vmatrix} + \vec{v}_1 \begin{vmatrix} a_{0,0} & a_{0,2} & \dots & a_{0,n} \\ a_{1,0} & a_{1,2} & \dots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,2} & \dots & a_{n-1,n} \end{vmatrix} + \dots$$

$$+ \vec{v}_n \begin{vmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n-1} \\ a_{1,0} & a_{1,1} & \dots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \dots & a_{n-1,n-1} \end{vmatrix}.$$

The most familiar example of a formal determinant is the three-dimensional cross product:

$$\vec{A} \times \vec{B} = \begin{vmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

The basic feature of formal determinants is that their inner product with another vector becomes a true determinant:

$$\left\langle \begin{vmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n} \\ a_{1,0} & a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \dots & a_{n-1,n} \\ \vec{v}_0 & \vec{v}_1 & \dots & \vec{v}_n \end{vmatrix}, w \right\rangle = \begin{vmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n} \\ a_{1,0} & a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \dots & a_{n-1,n} \\ \langle \vec{v}_0, w \rangle & \langle \vec{v}_1, w \rangle & \dots & \langle \vec{v}_n, w \rangle \end{vmatrix}$$

This may be easily verified by expanding both sides in cofactors along the bottom row.

In the case where $\vec{v}_0, \ldots, \vec{v}_n$ are orthonormal, the inner product of two formal determinants with the same row of vectors has a nice form:

$$\left\langle \begin{vmatrix} a_{0,0} & a_{0,1} & \ldots & a_{0,n} \\ a_{1,0} & a_{1,1} & \ldots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \ldots & a_{n-1,n} \\ \vec{v}_0 & \vec{v}_1 & \ldots & \vec{v}_n \end{vmatrix}, \begin{vmatrix} b_{0,0} & b_{0,1} & \ldots & b_{0,n} \\ b_{1,0} & b_{1,1} & \ldots & b_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n-1,0} & b_{n-1,1} & \ldots & b_{n-1,n} \\ \vec{v}_0 & \vec{v}_1 & \ldots & \vec{v}_n \end{vmatrix} \right\rangle$$
$$= \begin{vmatrix} \mathbf{a_0} \cdot \mathbf{b_0} & \mathbf{a_0} \cdot \mathbf{b_1} & \ldots & \mathbf{a_0} \cdot \mathbf{b_{n-1}} \\ \mathbf{a_1} \cdot \mathbf{b_0} & \mathbf{a_1} \cdot \mathbf{b_1} & \ldots & \mathbf{a_1} \cdot \mathbf{b_{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a_{n-1}} \cdot \mathbf{b_0} & \mathbf{a_{n-1}} \cdot \mathbf{b_1} & \ldots & \mathbf{a_{n-1}} \cdot \mathbf{b_{n-1}} \end{vmatrix}$$

where $\mathbf{a_k} = (a_{k,0}, \ldots, a_{k,n})$ and $\mathbf{b_k} = (b_{k,0} \ldots, b_{k,n})$ are the rows of the formal determinant viewed as vectors in F^{n+1}. (We use boldface instead of arrow and $\cdot$ instead of $\langle , \rangle$ to emphasize that this is in general a different space from V.) This may be proved, for example, by noting that both sides of the equation are linear in each of the $\mathbf{a_k}$'s and $\mathbf{b_k}$'s and verifying it in the case where all the $\mathbf{a_k}$'s and $\mathbf{b_k}$'s are unit coordinate vectors.

B.3 Gram-Schmidt Formula

We now can easily verify (B.1). First we note that the independence of the $\vec{v}_k$'s implies that $D_n > 0$, so that the $\vec{e}_k$'s are well-defined [Greub]. We will define the vectors $\vec{e}_k$ as in (B.1) and show that they satisfy the desired properties. For $j < k$, we have

$$\sqrt{D_{k-1}D_k}\langle \vec{v}_j, \vec{e}_k \rangle = \begin{vmatrix} \langle \vec{v}_0, \vec{v}_0 \rangle & \langle \vec{v}_1, \vec{v}_0 \rangle & \ldots & \langle \vec{v}_k, \vec{v}_0 \rangle \\ \langle \vec{v}_0, \vec{v}_1 \rangle & \langle \vec{v}_1, \vec{v}_1 \rangle & \ldots & \langle \vec{v}_k, \vec{v}_1 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{v}_0, \vec{v}_{k-1} \rangle & \langle \vec{v}_1, \vec{v}_{k-1} \rangle & \ldots & \langle \vec{v}_k, \vec{v}_{k-1} \rangle \\ \langle \vec{v}_0, \vec{v}_j \rangle & \langle \vec{v}_1, \vec{v}_j \rangle & \ldots & \langle \vec{v}_k, \vec{v}_j \rangle \end{vmatrix} = 0$$

because two rows in the determinant are identical. Because $\vec{e}_j$ is a linear combination of $\vec{v}_0, \ldots, \vec{v}_j$, it follows that $\langle \vec{e}_j, \vec{e}_k \rangle = 0$. Now because $\vec{e}_k$ is a linear combination of $\vec{v}_0, \ldots, \vec{v}_k$, and since the $\vec{e}$'s are orthogonal, we have $\operatorname{span}(\vec{e}_0, \ldots, \vec{e}_k) = \operatorname{span}(\vec{v}_0, \ldots, \vec{v}_k)$. Now if we expand $\vec{e}_k$, the coefficient of

$\vec{v}_k$ will be $\sqrt{\frac{D_{k-1}}{D_k}}$, so we have $\vec{e}_k = \sqrt{\frac{D_{k-1}}{D_k}}\vec{v}_k + \vec{w}_k$ where $\vec{w}_k$ is orthogonal to $\vec{e}_k$. Then $\langle \vec{e}_k, \vec{e}_k \rangle = \sqrt{\frac{D_{k-1}}{D_k}}\langle \vec{e}_k, \vec{v}_k \rangle = \sqrt{\frac{D_{k-1}}{D_k}}\left(\frac{1}{\sqrt{D_{k-1}D_k}}D_k\right) = 1.$

Appendix C

H^p Spaces on Polydiscs

In this appendix I will mention some of the more relevant facts from the theory of H^p spaces, first in the one-variable case and then in the more general setting of polydiscs. For details and proofs the reader is referred to [Duren] and [Rudin (1969)].

C.1 H^p Spaces in One Variable

Definition 11. *Let D denote the open unit disk in the complex plane. For an analytic function $f : D \to \mathbb{C}$, we let*

$$\begin{aligned} M_p(r,f) &= \left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(re^{i\theta})|^p \, d\theta\right\}^{1/p} \qquad 0 < p < \infty \\ M_\infty(r,f) &= \max_{-\pi \le \theta < \pi} |f(re^{i\theta})|. \end{aligned}$$

Note that for $1 \le p \le \infty$, $M(r,f) = \|f_r\|_p$ where $f_r : T \to \mathbb{C}$ is defined by $f_r(\theta) = f(re^{i\theta})$ and $\|\|_p$ is the L^p norm with respect to normalized Lebesgue measure on T. We can analogously define the quantities $M_p(r,u)$ for a harmonic function $u : D \to \mathbb{R}$.

Definition 12. *For $0 < p \le \infty$, H^p is the set of all analytic $f : D \to \mathbb{C}$ for which $M_p(r,f)$ is bounded as $r \to 1$. Similarly, h^p is the set of all harmonic $u : D \to \mathbb{R}$ for which $M_p(r,u)$ is bounded as $r \to 1$.*

The notation is in honor of Hardy, who was the first to study these spaces systematically.

As is well known, for a continuous function u on T, the Poisson integral formula

$$u(z) = u(re^{i\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1-r^2}{1-2r\cos(\theta)+r^2} u(e^{it})\,dt$$

gives a function $u : \bar{D} \to \mathbb{R}$ that is continuous on $\bar{D}$, harmonic on D, and has a specified set of boundary values. For a measure μ on T with bounded total variation, we can define the Poisson-Stieltjes integral

$$u(z) = u(re^{i\theta}) = \int_{-\pi}^{\pi} \frac{1-r^2}{1-2r\cos(\theta)+r^2}\,d\mu$$

which is again harmonic on D (although not necessarily continuous on $\bar{D}$).

Theorem 28. *The following three classes of functions from D to $\mathbb{R}$ are identical:*

- *Poisson-Stieltjes integrals;*
- *differences of two positive harmonic functions;*
- *h^1.*

Theorem 29. *Let $u(z)$ be a Poisson-Stieltjes integral. If the symmetric derivative*

$$D\mu(\theta_0) = \lim_{t\to 0} \frac{\mu((\theta_0 - t, \theta_0 + t))}{2t}$$

exists at a point θ_0, then the radial limit $\lim_{r\to 1} u(re^{i\theta_0})$ exists and has the value $2\pi D\mu(\theta_0)$.

Corollary 4. *Each function $u \in h^1$, and thus every $f \in H^1$, has a radial limit almost everywhere.*

We will frequently use $f(e^{i\theta})$ to denote the almost everywhere existing radial limit of an H^1 function $f : D \to \mathbb{C}$.

Corollary 5. *If u is the Poisson integral of a function $\phi \in L^1(T)$, then $u(re^{i\theta}) \to \phi(\theta)$ almost everywhere.*

Theorem 30 (Hardy's convexity theorem)**.** *Let $f(z)$ be analytic in D, and let $0 < p \leq \infty$. Then*

- *$M_p(r, f)$ is a nondecreasing function of r;*
- *$\log M_p(r, f)$ is a convex function of $\log r$.*

This theorem implies (among many other things) that we could equivalently define H^p as the space of functions for which $M_p(r, f)$ has a finite limit as $r \to 1$. This limit is called the H^p norm of f.

We already know that radial limits exist almost everywhere for h^1 and H^1 functions. Much more is true, as the follow theorem indicates.

Theorem 31. *If $f \in H^p$ for $0 < p \leq \infty$, then the radial limit $f(e^{i\theta})$ exists almost everywhere, and $f(e^{i\theta}) \in L^p(T)$.*

Now that we know that an H^p function has an L^p boundary function, the natural question is whether the H^p norm equals the L^p norm of the boundary. The following theorem answers this in the affirmative for $p < \infty$.

Theorem 32. *If $f \in H^p$ with $0 < p < \infty$, then*

$$\lim_{r \to 1} \int_{-\pi}^{\pi} |f(re^{i\theta})^p \, d\theta = \int_{-\pi}^{\pi} |f(e^{i\theta})|^p \, d\theta$$

and

$$\lim_{r \to 1} \int_{-\pi}^{\pi} |f(re^{i\theta}) - f(e^{i\theta})|^p \, d\theta = 0.$$

Of particular importance to our study of Szegö's theorem is the role of the Poisson integral in the interaction between functions on D and their boundary values. The next theorem explores this further.

Theorem 33. *For $1 \leq p \leq \infty$, an analytic function $f : D \to \mathbb{C}$ is the Poisson integral of an L^p function $\phi : T \to \mathbb{C}$ if and only if $f \in H^p$.*

It is now convenient to introduce the notation $\mathcal{H}^p$ to denote the set of boundary functions $f(e^{i\theta})$ of functions $f \in H^p$. (As with L^p, $\mathcal{H}^p$ is actually a space of equivalence classes, since the boundary values are only defined almost everywhere.) We know from the previous theorem that $\mathcal{H}^p \subset L^p(T)$, and it is evidently a linear subspace of $L^p(T)$. It turns out that $\mathcal{H}^p$ is topologically closed as well.

Theorem 34. *For $1 < p < \infty$, $\mathcal{H}^p$ is the L^p closure of the set of polynomials in $e^{i\theta}$.*

It can be shown that $\mathcal{H}^\infty$ is closed as well.

Corollary 6. *H^p is a Banach space for $1 \leq p \leq \infty$.*

The following theorem provides a simple characterization of $\mathcal{H}^p$.

Theorem 35. *For $1 \le p \le \infty$, $\mathcal{H}^p$ is exactly the set of L^p functions whose negative Fourier coefficients vanish. Moreover, if $f(z) = \sum_{n=0}^{\infty} a_n z^n \in H^1$, then $c_n = a_n$ for $n \ge 0$, where c_n is the nth Fourier coefficient of the boundary function $f(e^{i\theta})$.*

Note that among other things, this theorem implies that different functions in H^p have different boundary functions, so that there is a one-to-one correspondence between H^p and $\mathcal{H}^p$. The nature of this correspondence is further illuminated by the following theorem.

Theorem 36. *If $f \in L^p(T)$ and $P[f]$ is the Poisson integral of f, then $P[f]$ has f as its radial limit almost everywhere.*

We can now give an exact description of the correspondence between H^p and $\mathcal{H}^p$. Given a function $f \in H^p$, we know from Theorem 31 that the radial limit $f(e^{i\theta})$ exists almost everywhere; this defines (a.e.) a function $\tilde{f} : T \to \mathbb{C}$ which, by definition, is in $\mathcal{H}^p$. On the other hand, suppose we are given a function $\tilde{f} : T \to \mathbb{C}$ which is in $L^p(T)$ and whose negative Fourier coefficients vanish. Theorem 35 tells us that there exists a unique $f \in H^p$ that has $\tilde{f}$ as its radial limit a.e. Then by Theorem 33, f must be the Poisson integral of some $\phi \in L^p(T)$, and by Theorem 36, $\phi = \tilde{f}$ a.e. To sum up, given $f \in H^p$ we can find the corresponding $\tilde{f} \in \mathcal{H}^p$ by taking the radial limits of f; given $\tilde{f} \in \mathcal{H}^p$ we can find the corresponding $f \in H^p$ by taking the Poisson integral of $\tilde{f}$.

Because of this duality, most books do not explicitly mention $\mathcal{H}^p$ at all; it is considered to be identical with H^p. Thus, H^p is often spoken of as a subspace of $L^p(T)$, by identification with $\mathcal{H}^p \subset L^p(T)$.

Although H^p spaces are interesting in themselves, the following theorem is what connects them to our study of Toeplitz operators.

Theorem 37. *Let $g \in C(T)$ with $\int g d\theta = 0$. Then $e^g - 1 \in C(T)$ and $\int (e^g - 1) d\theta = 0$.*

Proof. Note that $g \in \mathcal{H}^\infty$. Let $G : D \to \mathbb{C}$ be the Poisson integral of g, so $G \in H^\infty$ by Theorem 33. Then $e^G - 1 \in H^\infty$ and is the Poisson integral of $e^g - 1$ (since its radial limit is $e^g - 1$ a.e. and H^p functions are Poisson integrals of their radial limits). Now

$$G(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\theta) d\theta = 0 \Rightarrow 0 = e^{G(0)} - 1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (e^{g(\theta)} - 1) d\theta$$

by elementary properties of the Poisson integral. □

C.2 H^p Spaces and Poisson Integrals on Polydiscs

A polydisc is a Cartesian product of several discs in the complex plane; in particular, $\mathbb{C}^n \supset D^n = \{(z_1, \ldots, z_n) \mid z_i \in D \text{ for } i = 1, \ldots, n\}$. We will focus on D^n since any other polydisc is equivalent to it by a simple change of variables.

For $n > 1$, T^n is only a small part of the boundary of D^n (it is, after all, only n-dimensional); however, for our purposes it is the important part. We will use $\overline{D}^n$ to denote the closure of the product of D's, not the product of the closures (so it includes the "unimportant" part of the boundary).

A continuous function $f : V \to \mathbb{C}$ on some open set $V \subset \mathbb{C}^n$ is said to be holomorphic if it is holomorphic in each variable separately. It is said to be n-harmonic if it is harmonic in each variable separately. Note that holomorphic functions are n-harmonic.

For a function $f : D^n \to \mathbb{C}$, we define $f_r : T^n \to \mathbb{C}$ by $f_r(\mathbf{w}) = f(r\mathbf{w})$. We use m_n to denote normalized Lebesgue measure on T^n; L^p spaces will be with respect to m_n unless otherwise noted.

The definition of the Hardy spaces is analogous to the one-variable case.

Definition 13. *For $0 < p \leq \infty$, $H^p(D^n)$ is the set of holomorphic $f : D^n \to \mathbb{C}$ for which*

$$\int_{T^n} |f_r(\mathbf{w})|^p dm_n(\mathbf{w})$$

is bounded for $0 \leq r < 1$, where $\mathbf{w} \in T^n$. The corresponding set of n-harmonic functions is denoted by $h^p(T^n)$.

Much of the theory of H^p spaces and Poisson integrals in one variable carries over very straightforwardly to polydiscs. In what follows we recount the most relevant results.

Definition 14. *If $\mathbf{z} \in D^n$, $\mathbf{w} \in T^n$, $z_j = r_j e^{i\theta_j}$, $w_j = e^{i\phi_j}$, the Poisson kernel $P(\mathbf{z}, \mathbf{w})$ is the product*

$$P(\mathbf{z}, \mathbf{w}) = P_{r_1}(\theta_1 - \phi_1) \ldots P_{r_n}(\theta_n - \phi_n)$$

where $P_r(\theta) = \frac{1-r^2}{1-2r\cos(\theta)+r^2}$ is the one-variable Poisson kernel.

Note that we have

$$\int_{T^n} P(\mathbf{z}, \mathbf{w}) dm_n(\mathbf{w}) = 1$$

by a simple application of Fubini's theorem, and that

$$P(\mathbf{z},\mathbf{w}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} r_1^{|k_1|}\dots r_n^{|k_n|} e^{i\mathbf{k}\cdot(\boldsymbol{\theta}-\boldsymbol{\phi})}$$

by multiplying the absolutely convergent series expansions in each variable individually. Here we use the notation $\mathbf{k} = (k_1,\dots,k_n) \in \mathbb{Z}^n$ and $\mathbf{k}\cdot(\boldsymbol{\theta} - \boldsymbol{\phi}) = k_1(\theta_1-\phi_1)+\cdots+k_n(\theta_n-\phi_n)$.

Definition 15. *For a complex Borel measure μ on T^n, the Poisson integral is*

$$P[d\mu](\mathbf{z}) = \int_{T^n} P(\mathbf{z},\mathbf{w})d\mu(\mathbf{w}) \qquad \textit{for } \mathbf{z}\in D^n.$$

For $f \in L^1(T^n)$, we let $P[f]$ denote $P[f\,dm_n]$.

Replacing $P(\mathbf{z},\mathbf{w})$ with its (uniformly convergent) series expansion and integrating termwise yields

$$P[d\mu](\mathbf{z}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} \hat{\mu}(\mathbf{k}) r_1^{|k_1|}\dots r_n^{|k_n|} e^{i\mathbf{k}\cdot\boldsymbol{\theta}}$$

where

$$\hat{\mu}(\mathbf{k}) = \int_{T^n} \overline{\mathbf{w}}^{\mathbf{k}} d\mu(\mathbf{w})$$

are the Fourier coefficients of μ. (Here $\overline{\mathbf{w}}^{\mathbf{k}}$ denotes $\overline{w_1}^{k_1}\dots\overline{w_n}^{k_n}$.)

Theorem 38. *If $u : \overline{D}^n \to \mathbb{C}$ is continuous and is n-harmonic in D^n, then $u(\mathbf{z}) = P[u](\mathbf{z})$ for $\mathbf{z}\in D^n$.*

Theorem 39.
1. *If $f \in L^\infty(T^n)$ and $\mathbf{z}\in D^n$ then $|P[f](\mathbf{z})| \le \|f\|_\infty$. Equality for any $z \in D^n$ implies that f is constant a.e. on T^n.*
2. *If $f \in C(T^n)$ then $P[f]$ extends to a continuous function on $\overline{D}^n$.*
3. *If $1 \le p < \infty$, $f \in L^p(T^n)$, and $u = P[f]$, then $\|u_r\|_p \le \|f\|_p$ and $\|u_r - f\|_p \to 0$ as $r \to 1$.*
4. *If $u = P[d\mu]$ then $\|u_r\|_1 \le \|\mu\|$.*
5. *If $u \in h^1(T^n)$, there is a unique measure μ on T^n such that $u = P[d\mu]$.*

Theorem 40. *$P[d\mu]$ is holomorphic in D^n iff $\hat{\mu}(\mathbf{k}) = 0$ for all $\mathbf{k}$ outside $\mathbf{Z}^n_+$. (Here $\mathbf{Z}^n_+$ denotes the set of ordered n-tuples of nonnegative integers.)*

The next theorem will concern the polydisc algebras $A(D^n)$, which are defined as in the one-variable case.

Definition 16. *$A(D^n)$ is the set of continuous functions from $\overline{D}^n$ to $\mathbb{C}$ whose restriction to D^n is holomorphic.*

Theorem 41. *A function $g \in C(T^n)$ is the restriction of a member of $A(D^n)$ iff $\hat{g}(\mathbf{k}) = 0$ outside $\mathbb{Z}_+^n$.*

For a function u on D^n, we define a function u^* on a subset of T^n by

$$u^*(\mathbf{w}) = \lim_{r \to 1} u(r\mathbf{w})$$

at every $\mathbf{w}$ for which the limit exists.

Theorem 42. *If μ is a measure on T^n with derivative $f \in L^1(T^n)$ with respect to dm_n, and $u = P[d\mu]$, then $u^*(\mathbf{w}) = f(\mathbf{w})$ for almost all (wrt dm_n) $\mathbf{w} \in T^n$.*

Having developed this theory of functions on polydiscs, we can prove an analogy to Theorem 37 for polydiscs.

Theorem 43. *Let $g \in C(T^n)$ with $\hat{g}(\mathbf{k}) = 0$ outside $\mathbb{Z}_+^n$ and $\int_{T^n} g\, dm_n = 0$. Then $e^g - 1 \in C(T^n)$ with $\widehat{e^g}(\mathbf{k}) = 0$ outside $\mathbb{Z}_+^n$ and $\int_{T^n}(e^g - 1)\, dm_n = 0$.*

Proof. By Theorem 41, g is the restriction of a function $G \in A(D^n)$. Then $G(\mathbf{z}) = P[g](\mathbf{z})$ for all $\mathbf{z} \in D^n$ by Theorem 38. In particular,

$$G(0) = \int_{T^n} g\, dm_n = 0.$$

Now e^G is clearly also in $A(D^n)$; it is therefore also the Poisson integral of its boundary, and

$$0 = e^{G(0)} - 1 = \int_{T^n} (e^g - 1)\, dm_n$$

as claimed. $\square$

Bibliography

Dunford, N. and Schwartz, J. T. (1957). *Linear Operators*. Wiley Interscience, New York. This comprehensive tome is the "bible" on linear operator theory. General analysis topics like measure and integration, the Baire theorem and its consequences, and weak topologies are covered, followed by an in-depth study of different types of linear spaces and operators on them. The entire second volume is devoted to Hermitian operators on Hilbert spaces, including differential operators, Hilbert-Schmidt operators, and compact operators. Spectral theorems are given much attention throughout.

Duren, P. (1970). *Theory of H^p Spaces*. Academic Press, New York. This book develops in detail the theory of H^p spaces, which are spaces of analytic function on the open unit disc with a particular kind of nice behavior at the boundary. The theory from this book is used in an elegant proof of Szegö's theorem.

Grenander, U. and Szegö, G. (1958). *Toeplitz Forms and Their Applications*. University of California Press, Berkeley. The theorem that is the focus of my research is treated in detail for the special case of the real line, and some mention is made of how to generalize to other measure spaces. Also contains a large section on applications to probability and statistics and analytic functions.

Greub, W. (1963). *Linear Algebra*. Springer-Verlag, New York. This book covers advanced linear algebra topics, including tensor algebra, exterior algebra, dual spaces, determinant theory, and affine and unitary spaces. The material on the Gram determinant, which gives a closed form for Gram-Schmidt orthogonalization, was particularly helpful.

Hewitt, E. and Stromberg, K. (1965). *Real and Abstract Analysis*. Springer-Verlag, New York. This is one of the standard textbooks for measure theory. Covers Lebesgue integration on abstract measure spaces, Banach and Hilbert spaces, differentiation of functions and measures, and integration on product spaces.

Hoffman, K. (1961). *Banach Spaces of Analytic Functions*. Dover, New York. This book applies methods of functional analysis to the study of analytic functions. The focus is on H^p spaces; since these are Banach, theorems from functional analysis can be used to prove facts about Poisson integrals, harmonic functions, and other complex analysis topics. This book contains a particularly elegant proof of Szegö's theorem.

Krieger, H. A. (1965). Toeplitz operators on locally compact abelian groups. *Journal of Mathematics and Mechanics*, 14:439–478. This is the article that my thesis is built on. It gives a partial generalization of Szegö's theorem to locally compact Abelian groups, proving that the discrete part of a measure has no effect on the asymptotic eigenvalue distribution of the Toeplitz operators. It also explores several other generalizations of Toeplitz operators.

Lang, S. (1993). *Real and Functional Analysis*. Springer-Verlag, New York. This book contains a succinct treatment of a huge amount of analysis. Abstract integration, Banach and Hilbert spaces, differential calculus, Gelfand transforms, compact and Fredholm operators, differential forms, manifolds, and geometric measure theory are all covered in less than 600 pages.

Munkres, J. (2000). *Topology*. Prentice Hall, Upper Saddle River, NJ. General reference for introductory topology—covers set-theoretic topology, separability and countability axioms, and metrization theorems, as well as algebraic topology via fundamental groups.

Riesz, F. and Szölö-Nagy, B. (1955). *Functional Analysis*. Dover, New York. A classic mathematics textbook, this is a seminal work in functional analysis by one of its inventors. After covering measure theory and integration, it uses integral equations as an introduction to the general theory of Banach and Hilbert spaces and linear operators.

Rudin, W. (1969). *Function Theory in Polydiscs*. W.A. Benjamin, Inc., New York. This book describes how standard complex variable theory, especially relating to the Poisson integral formula and H^p spaces, can be extended to polydiscs. It is useful in extending Szegö's theorem to the torus.

Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw Hill, New York. Standard reference for introductory analysis—construction of the reals, metric topology, continuity and differentiation, Riemann integration, preview of measure theory.

Rudin, W. (1987). *Real and Complex Analysis*. McGraw Hill, New York. Covers measure theory and abstract integration, focusing on locally compact Hausdorff spaces and the Riesz representation theorem. Extensive applications to complex analysis are given. Also includes introductory material on Fourier transforms, Banach algebras, and H^p spaces.

Rudin, W. (1990). *Fourier Analysis on Groups*. Wiley Classics, New York. This book describes how Fourier analysis generalizes from the unit circle and the real line to locally compact Abelian groups. Basic theorems like the inversion theorem and Plancherel's theorem are extended to this setting. A closer look is then given to the interplay between the algebraic properties of an LCA group and the resulting analytic properties of the Fourier transform.

Rudin, W. (1991). *Functional Analysis*. McGraw Hill, New York. The first section covers the general theory of Banach and Hilbert spaces; this is done at an advanced level, assuming prior familiarity with the subject (e.g. from Rudin's *Real and Complex Analysis*). The second section develops the theory of the Fourier transform on R^n, with applications to differential equations and (surprisingly) the prime number theorem. The third section covers Banach algebras and spectral theory.